

To make sure the weight calculation process is correct. I tried several weight expressions methods, both using the statement 'weights =' and 'model.weights=', as well as 'sandwich.nuisance=TRUE'

Weight calculation expression 1

The expression 1 using 'weights='.

```

    icc=0.1;intercept=-1.4;vars=0.2;times=1

## data_gene5 is the data generation function

d1=data_gene5(k=50,m=50,icc=icc,intercept=intercept,seed=times,vars=vars)
d2=d1;d2$y=ifelse(d2$R==1,NA,d2$y)
d3=data.frame(y=d2$y,x=d2$x,cluster=d2$cluster,arm=d2$arm,missing=d2$R)

head(d3)

##      y      x cluster arm missing
## 1  1  0.2633466      1  0        0
## 2  1 -0.3085617      2  0        0
## 3  1 -0.6364105      3  0        0
## 4  1  0.2138994      4  0        0
## 5 NA  0.9150394      5  0        1
## 6  1 -0.1198645      6  0        0

# weight calculation
logs=glm(missing ~ x+arm, data = d3,
        family = binomial(link='logit'))
logs2=glmer(missing ~ x+arm+(1|cluster) , data = d3,
           family = binomial(link='logit'))

weight1=1/predict(logs,type="response")
weight2=1/predict(logs2,type="response")

print(weight1[1:10])

##      1      2      3      4      5      6      7
## 3.455527 5.222642 6.761704 3.573361 2.323914 4.531039 11.690063
##      8      9     10
## 3.414098 2.930279 5.836695

print(weight2[1:10])

##      1      2      3      4      5      6      7
## 2.588240 5.910393 10.869044 3.682869 2.303290 6.667285 9.600998
##      8      9     10
## 2.286268 2.675370 7.044173

#different methods have the same results. We can just use the previous one to
#calculate weight
weight11=1/expit(predict(logs))
weight22=1/expit(predict(logs2))

print(weight11[1:10])

##      1      2      3      4      5      6      7
## 3.455527 5.222642 6.761704 3.573361 2.323914 4.531039 11.690063

```

```
##           8           9           10
## 3.414098  2.930279  5.836695
```

```
print(weight22[1:10])
```

```
##           1           2           3           4           5           6           7
## 2.588240  5.910393 10.869044  3.682869  2.303290  6.667285  9.600998
##           8           9           10
## 2.286268  2.675370  7.044173
```

```
d3$weight=weight1 ## add weights to d3
d3$weight2=weight2
```

Example results:

```
# weight independent working correlation matrix
# without consideration of cluster effects
ipw_ind1=myTryCatch(geeDREstimation(formula=y~x+arm,
                                     id="cluster" , data = d3,
                                     nameMISS='missing',nameY='y',
                                     nameTRT='arm',
                                     weights = d3$weight,
                                     family = binomial("logit"),
                                     corstr = "independence"))
summary(ipw_ind1$value)
```

```
##           Estimates Model SE Robust SE   wald p
## (Intercept)    0.8724  0.02783   0.05054 17.26 0
## x              1.1480  0.03448   0.08029 14.30 0
## arm            1.4670  0.05492   0.10750 13.65 0
##
## Est. Correlation:  0
## Correlation Structure:  independence
## Est. Scale Parameter:  1.047
##
## Number of GEE iterations: 2
## Number of Clusters:  100   Maximum Cluster Size:  60
## Number of observations with nonzero weight:  3210
```

```
# weight independent working correlation matrix
# with consideration of cluster effects
ipw_clu_ind1=myTryCatch(geeDREstimation(formula=y~x+arm,
                                         id="cluster" , data = d3,
                                         nameMISS='missing',nameY='y',
                                         nameTRT='arm',
                                         weights = d3$weight2,
                                         family = binomial("logit"),
                                         corstr = "independence"))
summary(ipw_clu_ind1$value)
```

```
##           Estimates Model SE Robust SE   wald p
## (Intercept)    0.8626  0.02552   0.05309 16.25 0
## x              1.0640  0.03319   0.07880 13.51 0
## arm            1.4500  0.04937   0.12220 11.87 0
##
## Est. Correlation:  0
## Correlation Structure:  independence
```

```
## Est. Scale Parameter: 1.025
##
## Number of GEE iterations: 2
## Number of Clusters: 100 Maximum Cluster Size: 60
## Number of observations with nonzero weight: 3210

# weight exchangeable working correlation matrix
# without consideration of cluster effects
ipw_ex1=myTryCatch(geeDREstimation(formula=y~x+arm,
                                   id="cluster" , data = d3,
                                   nameMISS='missing',nameY='y',
                                   nameTRT='arm',
                                   weights = d3$weight,
                                   family = binomial("logit"),
                                   corstr = "exchangeable"))

summary(ipw_ex1$value)
```

```
##           Estimates Model SE Robust SE  wald p
## (Intercept)      0.869  0.02438   0.05038 17.25 0
## x                1.146  0.03415   0.07778 14.74 0
## arm              1.464  0.04898   0.10490 13.95 0
##
## Est. Correlation: -0.004654
## Correlation Structure: exchangeable
## Est. Scale Parameter: 1.043
##
## Number of GEE iterations: 4
## Number of Clusters: 100 Maximum Cluster Size: 60
## Number of observations with nonzero weight: 3210
```

```
# weight exchangeable working correlation matrix
# with consideration of cluster effects
ipw_clu_ex1=myTryCatch(geeDREstimation(formula=y~x+arm,
                                       id="cluster" , data = d3,
                                       nameMISS='missing',nameY='y',
                                       nameTRT='arm',
                                       weights = d3$weight2,
                                       family = binomial("logit"),
                                       corstr = "exchangeable"))

summary(ipw_clu_ex1$value)
```

```
##           Estimates Model SE Robust SE  wald p
## (Intercept)      0.862  0.02248   0.05342 16.14 0
## x                1.064  0.03295   0.07700 13.82 0
## arm              1.449  0.04427   0.12050 12.02 0
##
## Est. Correlation: -0.00451
## Correlation Structure: exchangeable
## Est. Scale Parameter: 1.024
##
## Number of GEE iterations: 4
## Number of Clusters: 100 Maximum Cluster Size: 60
## Number of observations with nonzero weight: 3210
```

The total results

True effects

[1] 1.331

Table 1: weight calculation expression 1 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.115	0.130	0.107	0.116	0.920	0.915	0
0.2	0.001	0.001	0.111	0.144	0.107	0.129	0.955	0.920	0
0.3	0.000	0.001	0.109	0.152	0.105	0.142	0.935	0.925	0
0.4	0.001	0.004	0.104	0.163	0.103	0.154	0.950	0.930	0
0.5	0.003	0.000	0.103	0.178	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.189	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.178	0.099	0.167	0.960	0.935	0
0.8	0.001	0.007	0.090	0.170	0.097	0.155	0.960	0.930	0
0.9	0.003	0.007	0.083	0.140	0.095	0.136	0.975	0.930	0

Table 2: weight calculation expression 1 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.114	0.130	0.107	0.115	0.920	0.915	0
0.2	0.002	0.001	0.110	0.144	0.106	0.128	0.945	0.925	0
0.3	0.000	0.000	0.109	0.152	0.104	0.141	0.940	0.925	0
0.4	0.001	0.004	0.104	0.164	0.103	0.154	0.955	0.930	0
0.5	0.003	0.001	0.102	0.179	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.190	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.179	0.099	0.167	0.955	0.935	0
0.8	0.001	0.007	0.090	0.171	0.097	0.155	0.965	0.930	0
0.9	0.002	0.007	0.083	0.140	0.094	0.136	0.975	0.930	0

Weight calculation expression 2

The weight calculation is same with the expression 1 but added statement ‘sandwich.nuisance=TRUE’

The results are similar with expression 1

The total results

True effects

[1] 1.331

Table 3: weight calculation expression 2 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.115	0.130	0.107	0.116	0.920	0.915	0
0.2	0.001	0.001	0.111	0.144	0.107	0.129	0.955	0.920	0
0.3	0.000	0.001	0.109	0.152	0.105	0.142	0.935	0.925	0
0.4	0.001	0.004	0.104	0.163	0.103	0.154	0.950	0.930	0
0.5	0.003	0.000	0.103	0.178	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.189	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.178	0.099	0.167	0.960	0.935	0
0.8	0.001	0.007	0.090	0.170	0.097	0.155	0.960	0.930	0
0.9	0.003	0.007	0.083	0.140	0.095	0.136	0.975	0.930	0

Table 4: weight calculation expression 2 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.114	0.130	0.107	0.115	0.920	0.915	0
0.2	0.002	0.001	0.110	0.144	0.106	0.128	0.945	0.925	0
0.3	0.000	0.000	0.109	0.152	0.104	0.141	0.940	0.925	0
0.4	0.001	0.004	0.104	0.164	0.103	0.154	0.955	0.930	0
0.5	0.003	0.001	0.102	0.179	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.190	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.179	0.099	0.167	0.955	0.935	0
0.8	0.001	0.007	0.090	0.171	0.097	0.155	0.965	0.930	0
0.9	0.002	0.007	0.083	0.140	0.094	0.136	0.975	0.930	0

Weight calculation expression 3

Expression 3 calculating weights by using ‘model.weights=’, also added ‘sandwich.nuisance=TRUE’.

To make sure the model statement is correct, I tried the example in the CRTgeeDR package

```
library('CRTgeeDR')
```

```
## CRTgeeDR package example:
```

```
data(data.sim)
```

```
head(data.sim)
```

```
##      IDPAT CLUSTER TRT      X1 JOB MARRIED      AGE HIV.KNOW RELIGION
## 1      1      1      1 -1.7016825 no      no 26.00003 13.36309 1.3703589
## 2      2      1      1 3.3650937 no      yes 27.67290 15.66084 0.1469835
## 3      3      1      1 6.3091946 no      no 23.52499 16.91014 -0.7147724
## 4      4      1      1 0.1050196 no      yes 14.47551 16.33865 1.3637287
## 5      5      1      1 3.2173205 no      no 30.11747 12.30813 -0.2501598
## 6      6      1      1 3.1544019 no      yes 28.13158 15.35363 -0.4017665
##      OUTCOME MISSING
## 1      NA      1
## 2      1      0
## 3      NA      1
## 4      1      0
## 5      NA      1
## 6      NA      1
```

```
print(paste('the missing percentage is:',sum(data.sim$MISSING)/dim(data.sim)[1]))
```

```
## [1] "the missing percentage is: 0.243"
```

We can see:

- when MISSING = 1 means the outcome value is missing
- when MISSING = 0 means the outcome value is observed.

The results without ‘sandwich.nuisance=TRUE’ statement

in this section, MISSING = 1 means missing and MISSING = 0 means observed.

```
ipwresults<-geeDREstimation(formula=OUTCOME~TRT,
                             id="CLUSTER", data = data.sim,
                             family = "binomial", corstr = "independence",
```

```

                                model.weights=I(MISSING==0)~TRT*AGE)
summary(ipwresults)

```

```

##              Estimates Model SE Robust SE   wald      p
## (Intercept)   0.6766  0.03101   0.07059 9.585 0.00e+00
## TRT           0.4620  0.04618   0.10450 4.421 9.83e-06
##
## Est. Correlation:  0
## Correlation Structure:  independence
## Est. Scale Parameter:  1.075
##
## Number of GEE iterations: 2
## Number of Clusters: 100   Maximum Cluster Size: 100
## Number of observations with nonzero weight: 7570

```

The results with 'sandwich.nuisance=TRUE' statement

```

ipwresults2<-geeDREstimation(formula=OUTCOME~TRT,
                             id="CLUSTER" , data = data.sim,
                             family = "binomial", corstr = "independence",
                             sandwich.nuisance=TRUE,
                             model.weights=I(MISSING==0)~TRT*AGE)
summary(ipwresults2)

```

```

##              Estimates Model SE Robust SE   wald      p
## (Intercept)   0.6766  0.03101   0.07059 9.585 0.00e+00
## TRT           0.4620  0.04618   0.10450 4.421 9.83e-06
##
## Est. Correlation:  0
## Correlation Structure:  independence
## Est. Scale Parameter:  1.075
##
## Number of GEE iterations: 2
## Number of Clusters: 100   Maximum Cluster Size: 100
## Number of observations with nonzero weight: 7570

```

We can see that those results are similar. Besides, to calculate the weights, the model.weights equals to $I(\text{Missing} = a) \sim \text{covariates}$. And missing = a means observed.

In my weighing calculation:

```

icc=0.1;intercept=-1.4;times=123;vars=0.2

d1=data_gene5(k=50,m=50,icc=icc,intercept=intercept,seed=times,vars=vars)
d2=d1;d2$y=ifelse(d2$R==1,NA,d2$y)
d3=data.frame(y=d2$y,x=d2$x,cluster=d2$cluster,arm=d2$arm,missing=d2$R)

print(paste('missing percentage',sum(is.na(d3$y))/dim(d3)[1]))

## [1] "missing percentage 0.300520208083233"
head(d3)

```

```

##   y          x cluster arm missing
## 1 1  0.45666337      1  0        0
## 2 1 -0.47370676      2  0        0
## 3 1 -0.39553656      3  0        0
## 4 0 -0.21080697      4  0        0

```

```
## 5 1 0.29213108      5 0      0
## 6 1 -0.07789836     6 0      0
```

Therefore, missing=0 means observed. The statement should be

```
geeDREstimation(formula=y~x+arm, nameTRT = "arm",nameMISS = "missing", nameY = "y",
id="cluster", data = d3, family = binomial("logit"), corstr = "exchangeable", sandwich.nuisance=TRUE,
model.weights=I(missing==0)~x+arm)
```

Example results:

```
icc=0.1;intercept=-1.4;times=1;vars=0.2
d1=data_gene5(k=50,m=50,icc=icc,intercept=intercept,seed=times,vars=vars)
d2=d1;d2$y=ifelse(d2$R==1,NA,d2$y)
d3=data.frame(y=d2$y,x=d2$x,cluster=d2$cluster,arm=d2$arm,missing=d2$R)

## independent working correlation matrix
## without cluster effects

ipw_ind3=myTryCatch(geeDREstimation(formula=y~x+arm,
                                   nameTRT = "arm",nameMISS = "missing", nameY = "y",
                                   id="cluster", data = d3,
                                   family = binomial("logit"),
                                   corstr = "independence",
                                   sandwich.nuisance=TRUE,
                                   model.weights=I(missing==0)~x+arm))

summary(ipw_ind3$value)
```

```
##              Estimates Model SE Robust SE   wald p
## (Intercept)    0.8545  0.04566   0.04960 17.23 0
## x              1.0420  0.06014   0.07038 14.80 0
## arm            1.3870  0.08139   0.09362 14.81 0
##
## Est. Correlation: 0
## Correlation Structure: independence
## Est. Scale Parameter: 0.9974
##
## Number of GEE iterations: 2
## Number of Clusters: 100   Maximum Cluster Size: 60
## Number of observations with nonzero weight: 3210
```

```
## independent working correlation matrix
## with cluster effects
ipw_clu_ind3=myTryCatch(geeDREstimation(formula=y~x+arm,
                                       id="cluster", data = d3,
                                       nameMISS='missing',nameY='y',
                                       nameTRT='arm',
                                       sandwich.nuisance=TRUE,
                                       model.weights=I(missing==0)~x+arm+(1|cluster),
                                       family = binomial("logit"),
                                       corstr = "independence"))

summary(ipw_clu_ind3$value)
```

```
##              Estimates Model SE Robust SE   wald p
## (Intercept)    0.8545  0.04566   0.04960 17.23 0
## x              1.0420  0.06014   0.07038 14.80 0
## arm            1.3870  0.08139   0.09362 14.81 0
```

```
##
## Est. Correlation: 0
## Correlation Structure: independence
## Est. Scale Parameter: 0.9974
##
## Number of GEE iterations: 2
## Number of Clusters: 100 Maximum Cluster Size: 60
## Number of observations with nonzero weight: 3210
## exchangeable working correlation matrix
## without cluster effects
ipw_ex3=myTryCatch(geeDREstimation(formula=y~x+arm,
                                   nameTRT = "arm",nameMISS = "missing", nameY = "y",
                                   id="cluster" , data = d3,
                                   family = binomial("logit"),
                                   corstr = "exchangeable",
                                   sandwich.nuisance=TRUE,
                                   model.weights=I(missing==0)~x+arm))
summary(ipw_ex3$value)
```

```
##           Estimates Model SE Robust SE   wald p
## (Intercept)    0.8507  0.03983   0.05009 16.98 0
## x              1.0430  0.05995   0.07060 14.77 0
## arm            1.3900  0.07136   0.09501 14.63 0
##
## Est. Correlation: -0.004738
## Correlation Structure: exchangeable
## Est. Scale Parameter: 0.9966
##
## Number of GEE iterations: 3
## Number of Clusters: 100 Maximum Cluster Size: 60
## Number of observations with nonzero weight: 3210
```

```
## exchangeable working correlation matrix
## with cluster effects
ipw_clu_ex3=myTryCatch(geeDREstimation(formula=y~x+arm,
                                       id="cluster" , data = d3,
                                       nameMISS='missing',nameY='y',
                                       nameTRT='arm',
                                       sandwich.nuisance=TRUE,
                                       model.weights=I(missing==0)~x+arm+(1|cluster),
                                       family = binomial("logit"),
                                       corstr = "exchangeable"))
summary(ipw_clu_ex3$value)
```

```
##           Estimates Model SE Robust SE   wald p
## (Intercept)    0.8507  0.03983   0.05009 16.98 0
## x              1.0430  0.05995   0.07060 14.77 0
## arm            1.3900  0.07136   0.09501 14.63 0
##
## Est. Correlation: -0.004738
## Correlation Structure: exchangeable
## Est. Scale Parameter: 0.9966
##
## Number of GEE iterations: 3
## Number of Clusters: 100 Maximum Cluster Size: 60
```


Number of observations with nonzero weight: 3210

The total results

True effects

[1] 1.331

Table 5: weight calculation expression 3 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.115	0.130	0.107	0.116	0.920	0.915	0
0.2	0.001	0.001	0.111	0.144	0.107	0.129	0.955	0.920	0
0.3	0.000	0.001	0.109	0.152	0.105	0.142	0.935	0.925	0
0.4	0.001	0.004	0.104	0.163	0.103	0.154	0.950	0.930	0
0.5	0.003	0.000	0.103	0.178	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.189	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.178	0.099	0.167	0.960	0.935	0
0.8	0.001	0.007	0.090	0.170	0.097	0.155	0.960	0.930	0
0.9	0.003	0.007	0.083	0.140	0.095	0.136	0.975	0.930	0

Table 6: weight calculation expression 3 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.114	0.130	0.107	0.115	0.920	0.915	0
0.2	0.002	0.001	0.110	0.144	0.106	0.128	0.945	0.925	0
0.3	0.000	0.000	0.109	0.152	0.104	0.141	0.940	0.925	0
0.4	0.001	0.004	0.104	0.164	0.103	0.154	0.955	0.930	0
0.5	0.003	0.001	0.102	0.179	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.190	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.179	0.099	0.167	0.955	0.935	0
0.8	0.001	0.007	0.090	0.171	0.097	0.155	0.965	0.930	0
0.9	0.002	0.007	0.083	0.140	0.094	0.136	0.975	0.930	0

Weight calculation expression 4

According to the package example, we should use the statement: ‘model.weights=I(missing==0)~x+arm’

However, when we compare the results with expression 1 and expression 2, we can find that:

```
icc=0.1;intercept=-1.4;vars=0.2;times=1
d1=data_gene5(k=50,m=50,icc=icc,intercept=intercept,seed=times,vars=vars)
d2=d1;d2$y=ifelse(d2$R==1,NA,d2$y)
d3=data.frame(y=d2$y,x=d2$x,cluster=d2$cluster,arm=d2$arm,missing=d2$R)

logs=glm(missing ~ x+arm, data = d3,
         family = binomial(link='logit'))

weight1=1/predict(logs,type="response")
d3$weight=weight1

r1=geeDREstimation(formula=y~x+arm,
                   id="cluster" , data = d3,
                   nameMISS='missing',nameY='y',
```

```

        nameTRT='arm',
        sandwich.nuisance=TRUE,
        weights=d3$weight,
        family = binomial("logit"),
        corstr = "independence")

r2=geeDREstimation(formula=y~x+arm,
                    id="cluster" , data = d3,
                    nameMISS='missing',nameY='y',
                    nameTRT='arm',
                    sandwich.nuisance=TRUE,
                    model.weights=I(missing==0)~arm+x,
                    family = binomial("logit"),
                    corstr = "independence")

r3=geeDREstimation(formula=y~x+arm,
                    id="cluster" , data = d3,
                    nameMISS='missing',nameY='y',
                    nameTRT='arm',
                    sandwich.nuisance=TRUE,
                    model.weights=I(missing==1)~arm+x,
                    family = binomial("logit"),
                    corstr = "independence")

summary(r1)

##              Estimates Model SE Robust SE   wald p
## (Intercept)    0.8724  0.02783   0.05054 17.26 0
## x              1.1480  0.03448   0.08029 14.30 0
## arm            1.4670  0.05492   0.10750 13.65 0
##
## Est. Correlation:  0
## Correlation Structure:  independence
## Est. Scale Parameter:  1.047
##
## Number of GEE iterations: 2
## Number of Clusters:  100   Maximum Cluster Size:  60
## Number of observations with nonzero weight:  3210

summary(r2)

##              Estimates Model SE Robust SE   wald p
## (Intercept)    0.8545  0.04566   0.04960 17.23 0
## x              1.0420  0.06014   0.07038 14.80 0
## arm            1.3870  0.08139   0.09362 14.81 0
##
## Est. Correlation:  0
## Correlation Structure:  independence
## Est. Scale Parameter:  0.9974
##
## Number of GEE iterations: 2
## Number of Clusters:  100   Maximum Cluster Size:  60
## Number of observations with nonzero weight:  3210

```

```
summary(r3)
```

```
##              Estimates Model SE Robust SE  wald p
## (Intercept)   0.8724  0.02783   0.05054 17.26 0
## x             1.1480  0.03448   0.08029 14.30 0
## arm           1.4670  0.05492   0.10750 13.65 0
##
## Est. Correlation:  0
## Correlation Structure:  independence
## Est. Scale Parameter:  1.047
##
## Number of GEE iterations: 2
## Number of Clusters:  100   Maximum Cluster Size:  60
## Number of observations with nonzero weight:  3210
```

We can find ‘model.weights=I(missing==1)~x+arm’ matches the results by using weights=d3\$weight

I don’t know why it is. Is there something wrong?

And I also ran the results for ‘model.weights=I(missing==1)~x+arm’

Total results

True effects

```
## [1] 1.331
```

Table 7: weight calculation expression 4 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.008	0.115	0.115	0.107	0.107	0.920	0.920	0
0.2	0.001	0.001	0.111	0.111	0.107	0.107	0.955	0.955	0
0.3	0.000	0.000	0.109	0.109	0.105	0.105	0.935	0.935	0
0.4	0.001	0.001	0.104	0.104	0.103	0.103	0.950	0.950	0
0.5	0.003	0.003	0.103	0.103	0.102	0.102	0.930	0.930	0
0.6	0.002	0.002	0.098	0.098	0.101	0.101	0.970	0.970	0
0.7	0.002	0.002	0.095	0.095	0.099	0.099	0.960	0.960	0
0.8	0.001	0.001	0.090	0.090	0.097	0.097	0.960	0.960	0
0.9	0.003	0.003	0.083	0.083	0.095	0.095	0.975	0.975	0

Table 8: weight calculation expression 4 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.008	0.114	0.114	0.107	0.107	0.920	0.920	0
0.2	0.002	0.002	0.110	0.110	0.106	0.106	0.945	0.945	0
0.3	0.000	0.000	0.109	0.109	0.104	0.104	0.940	0.940	0
0.4	0.001	0.001	0.104	0.104	0.103	0.103	0.955	0.955	0
0.5	0.003	0.003	0.102	0.102	0.102	0.102	0.930	0.930	0
0.6	0.002	0.002	0.098	0.098	0.101	0.101	0.970	0.970	0
0.7	0.002	0.002	0.095	0.095	0.099	0.099	0.955	0.955	0
0.8	0.001	0.001	0.090	0.090	0.097	0.097	0.965	0.965	0
0.9	0.002	0.002	0.083	0.083	0.094	0.094	0.975	0.975	0

Besides, there are also some errors for the statement 'sandwich.nuisance=TRUE'

With different settings of input data set, there may be some errors generated by this statement.

For example:

```
## This scenario works well
```

```
icc=0.1;intercept=-1.4;times=1;vars=0.2
d1=data_gene5(k=50,m=50,icc=icc,intercept=intercept,seed=times,vars=vars)
d2=d1;d2$y=ifelse(d2$R==1,NA,d2$y)
d3=data.frame(y=d2$y,x=d2$x,cluster=d2$cluster,arm=d2$arm,missing=d2$R)
head(d3)
```

```
##      y      x cluster arm missing
## 1  1  0.2633466      1  0        0
## 2  1 -0.3085617      2  0        0
## 3  1 -0.6364105      3  0        0
## 4  1  0.2138994      4  0        0
## 5 NA  0.9150394      5  0        1
## 6  1 -0.1198645      6  0        0
```

```
geeDREstimation(formula=y~arm+x,
                 nameTRT = "arm",nameMISS = "missing", nameY = "y",
                 id="cluster" , data = d3,
                 family = binomial("logit"), corstr = "independence",
                 sandwich.nuisance=TRUE,
                 model.weights=I(missing==0)~arm+x)
```

```
## geeDREstimation(formula = y ~ arm + x, id = "cluster", data = d3,
##      family = binomial("logit"), corstr = "independence", nameTRT = "arm",
##      model.weights = I(missing == 0) ~ arm + x, nameMISS = "missing",
##      nameY = "y", sandwich.nuisance = TRUE)
##
## Coefficients:
## (Intercept)  arm      x
##      0.8545 1.387 1.042
##
## Scale Parameter:  0.9974
##
## Correlation Model:  independence
## Estimated Correlation Parameters:  0
##
## Number of clusters:  100   Maximum cluster size:  60
## Number of observations with nonzero weight:  3210
## However, an error generated
```

```
icc=0.9;intercept=-4;times=500;vars=0.2
d1=data_gene5(k=50,m=50,icc=icc,intercept=intercept,seed=times,vars=vars)
d2=d1;d2$y=ifelse(d2$R==1,NA,d2$y)
d3=data.frame(y=d2$y,x=d2$x,cluster=d2$cluster,arm=d2$arm,missing=d2$R)
head(d3)
```

```
##      y      x cluster arm missing
```

```
## 1 1 1.4790889      1 0      0
## 2 1 -0.4250301     2 0      0
## 3 0 0.2048202      3 0      0
## 4 1 0.8743999      4 0      0
## 5 1 0.1233705      5 0      0
## 6 1 0.3966916      6 0      0
```

```
res=myTryCatch(geeDREstimation(formula=y~arm+x,
                               nameTRT = "arm",nameMISS = "missing", nameY = "y",
                               id="cluster" , data = d3,
                               family = binomial("logit"), corstr = "independence",
                               sandwich.nuisance=TRUE,
                               model.weights=I(missing==0)~arm+x))
```

```
## There was an error in the nuisance variance computation
```

```
print(res$warn)
```

```
## <simpleWarning in as.numeric(diag(sqrtW) > 0)/(exp((design.weights %*% etaW))/(1 + exp((design.w
```

```
print(res$error)
```

```
## <simpleError in geeDREstimation(formula = y ~ arm + x, nameTRT = "arm", nameMISS = "missing", na
```

- The error is:

```
simpleError in geeDREstimation(formula = y ~ arm + x, nameTRT = "arm", nameMISS = "missing", nameY = "y", id = "cluster", data = d3, family = binomial("logit"), corstr = "independence", sandwich.nuisance = TRUE, model.weights = I(missing == 0) ~ arm + x): object 'sandvarnuis.list' not found
```

- The warning is:

```
simpleWarning in as.numeric(diag(sqrtW) > 0)/(exp((design.weights %*% etaW))/(1 + exp((design.weights %*% etaW)))): longer object length is not a multiple of shorter object length>
```

```
.
.
.
```

Draw the tables together to be more clear

True effects

```
## [1] 1.331
```

Table 1: weight calculation expression 1 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.115	0.130	0.107	0.116	0.920	0.915	0
0.2	0.001	0.001	0.111	0.144	0.107	0.129	0.955	0.920	0
0.3	0.000	0.001	0.109	0.152	0.105	0.142	0.935	0.925	0
0.4	0.001	0.004	0.104	0.163	0.103	0.154	0.950	0.930	0
0.5	0.003	0.000	0.103	0.178	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.189	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.178	0.099	0.167	0.960	0.935	0
0.8	0.001	0.007	0.090	0.170	0.097	0.155	0.960	0.930	0
0.9	0.003	0.007	0.083	0.140	0.095	0.136	0.975	0.930	0

Table 2: weight calculation expression 1 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.114	0.130	0.107	0.115	0.920	0.915	0
0.2	0.002	0.001	0.110	0.144	0.106	0.128	0.945	0.925	0
0.3	0.000	0.000	0.109	0.152	0.104	0.141	0.940	0.925	0
0.4	0.001	0.004	0.104	0.164	0.103	0.154	0.955	0.930	0
0.5	0.003	0.001	0.102	0.179	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.190	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.179	0.099	0.167	0.955	0.935	0
0.8	0.001	0.007	0.090	0.171	0.097	0.155	0.965	0.930	0
0.9	0.002	0.007	0.083	0.140	0.094	0.136	0.975	0.930	0

Table 3: weight calculation expression 2 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.115	0.130	0.107	0.116	0.920	0.915	0
0.2	0.001	0.001	0.111	0.144	0.107	0.129	0.955	0.920	0
0.3	0.000	0.001	0.109	0.152	0.105	0.142	0.935	0.925	0
0.4	0.001	0.004	0.104	0.163	0.103	0.154	0.950	0.930	0
0.5	0.003	0.000	0.103	0.178	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.189	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.178	0.099	0.167	0.960	0.935	0
0.8	0.001	0.007	0.090	0.170	0.097	0.155	0.960	0.930	0
0.9	0.003	0.007	0.083	0.140	0.095	0.136	0.975	0.930	0

Table 4: weight calculation expression 2 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.114	0.130	0.107	0.115	0.920	0.915	0
0.2	0.002	0.001	0.110	0.144	0.106	0.128	0.945	0.925	0
0.3	0.000	0.000	0.109	0.152	0.104	0.141	0.940	0.925	0
0.4	0.001	0.004	0.104	0.164	0.103	0.154	0.955	0.930	0
0.5	0.003	0.001	0.102	0.179	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.190	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.179	0.099	0.167	0.955	0.935	0
0.8	0.001	0.007	0.090	0.171	0.097	0.155	0.965	0.930	0
0.9	0.002	0.007	0.083	0.140	0.094	0.136	0.975	0.930	0

Table 5: weight calculation expression 3 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.115	0.130	0.107	0.116	0.920	0.915	0
0.2	0.001	0.001	0.111	0.144	0.107	0.129	0.955	0.920	0
0.3	0.000	0.001	0.109	0.152	0.105	0.142	0.935	0.925	0
0.4	0.001	0.004	0.104	0.163	0.103	0.154	0.950	0.930	0
0.5	0.003	0.000	0.103	0.178	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.189	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.178	0.099	0.167	0.960	0.935	0
0.8	0.001	0.007	0.090	0.170	0.097	0.155	0.960	0.930	0
0.9	0.003	0.007	0.083	0.140	0.095	0.136	0.975	0.930	0

Table 6: weight calculation expression 3 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.114	0.130	0.107	0.115	0.920	0.915	0
0.2	0.002	0.001	0.110	0.144	0.106	0.128	0.945	0.925	0
0.3	0.000	0.000	0.109	0.152	0.104	0.141	0.940	0.925	0
0.4	0.001	0.004	0.104	0.164	0.103	0.154	0.955	0.930	0
0.5	0.003	0.001	0.102	0.179	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.190	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.179	0.099	0.167	0.955	0.935	0
0.8	0.001	0.007	0.090	0.171	0.097	0.155	0.965	0.930	0
0.9	0.002	0.007	0.083	0.140	0.094	0.136	0.975	0.930	0

Table 7: weight calculation expression 4 with independent working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.115	0.130	0.107	0.116	0.920	0.915	0
0.2	0.001	0.001	0.111	0.144	0.107	0.129	0.955	0.920	0
0.3	0.000	0.001	0.109	0.152	0.105	0.142	0.935	0.925	0
0.4	0.001	0.004	0.104	0.163	0.103	0.154	0.950	0.930	0
0.5	0.003	0.000	0.103	0.178	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.189	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.178	0.099	0.167	0.960	0.935	0
0.8	0.001	0.007	0.090	0.170	0.097	0.155	0.960	0.930	0
0.9	0.003	0.007	0.083	0.140	0.095	0.136	0.975	0.930	0

Table 8: weight calculation expression 4 with exchangeable working correlation matrix

ICC	Bias		MCSD		SD		Coverage		non
	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	IPW	IPW_clu	
0.1	0.008	0.005	0.114	0.130	0.107	0.115	0.920	0.915	0
0.2	0.002	0.001	0.110	0.144	0.106	0.128	0.945	0.925	0
0.3	0.000	0.000	0.109	0.152	0.104	0.141	0.940	0.925	0
0.4	0.001	0.004	0.104	0.164	0.103	0.154	0.955	0.930	0
0.5	0.003	0.001	0.102	0.179	0.102	0.165	0.930	0.930	0
0.6	0.002	0.001	0.098	0.190	0.101	0.170	0.970	0.900	0
0.7	0.002	0.010	0.095	0.179	0.099	0.167	0.955	0.935	0
0.8	0.001	0.007	0.090	0.171	0.097	0.155	0.965	0.930	0
0.9	0.002	0.007	0.083	0.140	0.094	0.136	0.975	0.930	0