#####Set working directory (PLEASE REPLACE IT WITH YOUR OWN PATH)

setwd("D:/")

#####Load sample data

array<-read.table("Sample\_data.txt",header=T)

m = nrow(array)

n = ncol(array)-3

#####Simulate 10000 loci and calculate their scores

score<-thre<-skewness<-y\_mean<-numeric(1)

nper=10000

random\_control <- matrix(runif(nper\*n/2,0,1), nper, n/2)

random\_case <- matrix(runif(nper\*n/2,0,1), nper, n/2)

distance <- sqrt(random\_control^2 + random\_case^2)

y<-(random\_case - random\_control)\*(random\_case + random\_control)/(2\*distance)

for (i in 1:nper){

y\_mean[i]=mean(y[i,1:(n/2)])

sum=0

for(k in 1:(n/2)){

sum=sum+(y[i,k]-mean(y[i,1:(n/2)]))\*\*3

}

b1=sum\*n

b2=(n/2-1)\*(n/2-2)\*(sd(y[i,1:(n/2)])\*\*3)

skewness[i]=1/(1+exp(-abs(b1/b2)))

score[i] = abs(y\_mean[i]/skewness[i])

}

diff = sort(score)

thre = diff[nper\*(1-0.05)]

#####Identify DML

y\_mean<-distance\_x<-distance\_y<-score<-skewness<-PValue<-numeric(1)

a<-array[1:m,4:(n/2+3)]

b<-array[1:m,(n/2+4):(n+3)]

distance = sqrt(a\*a+b\*b)

y = (b-a)\*(b+a)/(2\*distance)

for (i in 1:m){

distance\_x[i]=mean(as.numeric(a[i,1:(n/2)]))

distance\_y[i]=mean(as.numeric(b[i,1:(n/2)]))

y\_mean[i]=mean(as.numeric(y[i,1:(n/2)]))

sum=0

for(k in 1:(n/2)){

sum=sum+(y[i,k]-mean(as.numeric(y[i,1:(n/2)])))\*\*3

}

b1=sum\*n

b2=(n/2-1)\*(n/2-2)\*(sd(as.numeric(y[i,1:(n/2)]))\*\*3)

skewness[i]=1/(1+exp(-abs(b1/b2)))

score[i] = y\_mean[i]/skewness[i]

GreaterNumbers=0;

for (f in 1:nper){

if (diff[f]>=abs(score[i])){

GreaterNumbers=GreaterNumbers+1

}

}

PValue[i]=GreaterNumbers/nper

}

#####Save output to your working directory

outdata = data.frame(array$ILMNID,array$CHR,distance\_x,distance\_y,skewness,score,PValue)

write.csv(outdata, file="csDML\_Output.csv", row.names=F)