Plotting\_Challenges.R

MaBe

Thu Sep 17 17:50:06 2015

#Exercise 1a  
setwd("/Users/MaBe/Desktop/Bootcamp\_Mac")  
snpsDataFrame=read.table('hapmap\_CEU\_r23a\_chr2\_ld-1.txt',header=TRUE)  
snps=as.matrix(snpsDataFrame)  
  
  
#Exercise 1b  
mean(pvals<0.05) #The proportion of P-values is 0.04509218

## [1] 0.04509218

mean(pvals<0.01) #The proportion of P-values is 0.01021425

## [1] 0.01021425

mean(pvals<0.001) #The proportion of P-values is 0.00124564

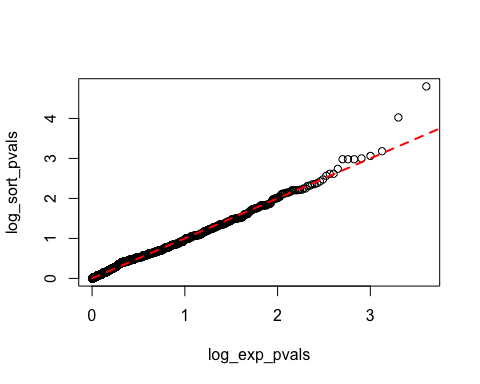
## [1] 0.00124564

#Exercise 1c  
length(pvals)

## [1] 4014

num\_pval<-length(pvals)  
  
  
#Exercise 1d  
nn<-(1:4014)  
exp\_pvals<-nn/num\_pval  
#I just created my expected P-values  
  
  
#Exercise 1e  
sort(pvals)

#Exercise 1g  
log\_sort\_pvals<--log10(sort\_pvals)  
plot(log\_exp\_pvals, log\_sort\_pvals)  
  
  
#Exercise 1h  
lines(0:4,0:4,col=2, lty=2, lwd=2)



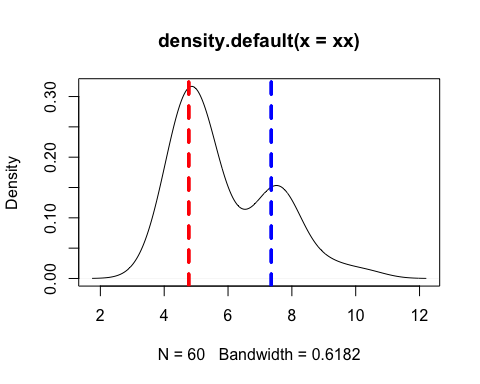
#I just add the line to show the relation between expected and observed values  
  
  
  
  
#Exercise 2a  
setwd("/Users/MaBe/Desktop/Bootcamp\_Mac")  
zz=read.table('pheno.sim.2014-1.txt',header=TRUE)  
  
  
#Exercise 2b  
xx<-zz[,2]  
quantile(xx) #The value of the phenotype where 25% of the individuals fall is 4.768756

## 0% 25% 50% 75% 100%   
## 3.595392 4.768756 5.327751 7.354975 10.357491

yy<-subset(zz,zz[,2]<4.768756) #I extracted IDs and phenotypes of the individuals fulfilling this criterion  
controls<-yy[,1] #I extracted IDs of the individuals fulfilling this criterion  
  
  
#Exercise 2c  
xx<-zz[,2]  
quantile(xx) #The value of the phenotype where 25% of the individuals fall is 7.354975

## 0% 25% 50% 75% 100%   
## 3.595392 4.768756 5.327751 7.354975 10.357491

ww<-subset(zz,zz[,2]>7.354975) #I extracted IDs and phenotypes of the individuals fulfilling this criterion  
case<-ww[,1] #I extracted IDs of the individuals fulfilling this criterion  
  
  
#Exercise 2d  
plot(density(xx))  
abline(v=quantile(4.768756),lty=2,lwd=3,col=2)  
abline(v=quantile(7.354975),lty=2,lwd=3,col=4)



#Exercise 2e  
caseSNP=snps["rs7584086\_T",]  
case\_genotypes<-caseSNP[case]  
  
  
#Exercise 2f  
caseSNP=snps["rs7584086\_T",]  
control\_genotypes<-caseSNP[controls]  
  
  
#Exercise 2g  
Tcase<-table(case\_genotypes)