HW4-Solution

STAT-COMP 2020

12/11/2020

Single marker regression

```
load('~/Dropbox/STATCOMP/2020/Xy.RData')

## Single-marker regression test
 SMR=matrix(nrow=ncol(X),ncol=4,NA)
 colnames(SMR)=c('Estimate','SE','z-stat','p-value')

 for(i in 1:ncol(X)){
  SMR[i,]=ls.print(lsfit(y=y,x=X[,i]),print.it=F)[[2]][[1]][2,]
 }

 head(SMR)
```

```
##          Estimate          SE     z-stat    p-value
## [1,] 0.066375838 0.11438827 0.58026785 0.5617471
## [2,] 0.071140176 0.12721343 0.55921907 0.5760247
## [3,] 0.019019788 0.05703887 0.33345310 0.7387993
## [4,] 0.007430585 0.05438721 0.13662376 0.8913309
## [5,] 0.002325077 0.17841395 0.01303193 0.9896026
## [6,] 0.021287279 0.05337528 0.39882278 0.6900323
```
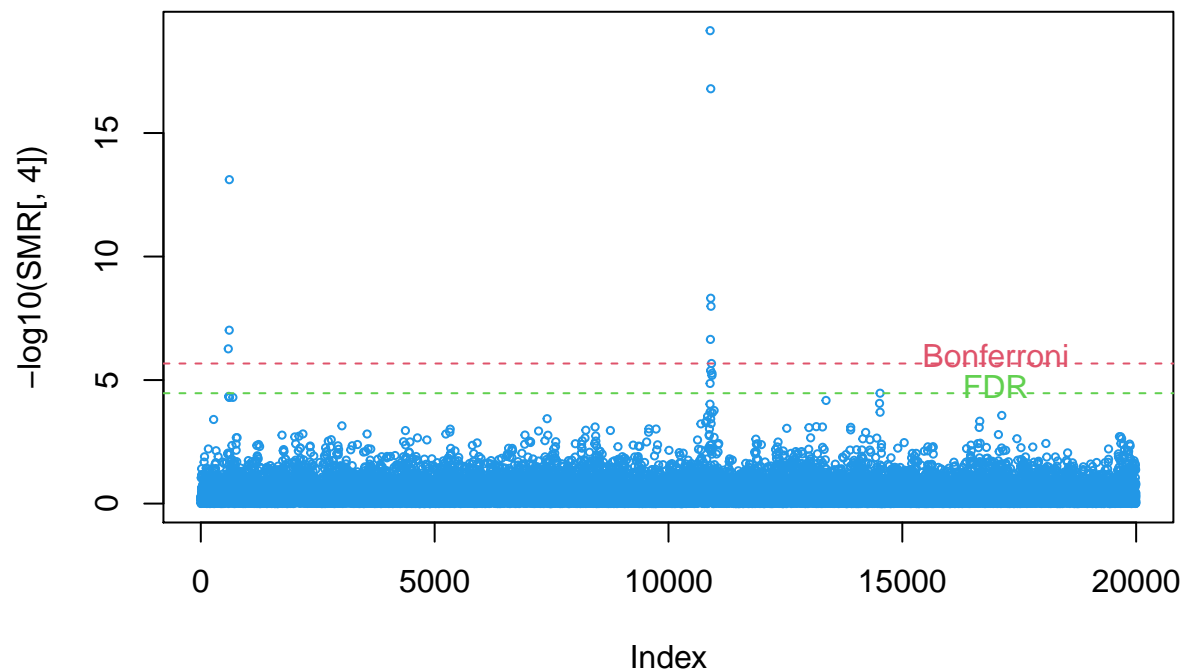
Adjusting p-values and determining significance

```
BONF=p.adjust(SMR[,4],method='bonferroni')
BONF_SIG=BONF<0.05

FDR=p.adjust(SMR[,4],method='fdr')
FDR_SIG=FDR<0.05
```

Manhattan plot

```
## Manhattan plot
 plot(-log10(SMR[,4]),cex=.5,col=4 )
 abline(h= -log10(max(SMR[BONF_SIG,4])),lty=2,col=2)
 abline(h= -log10(max(SMR[FDR_SIG,4])),lty=2,col=3)
 text(x=17000,y=5.98, label='Bonferroni',col=2)
 text(x=17000,y=4.75, label='FDR',col=3)
```
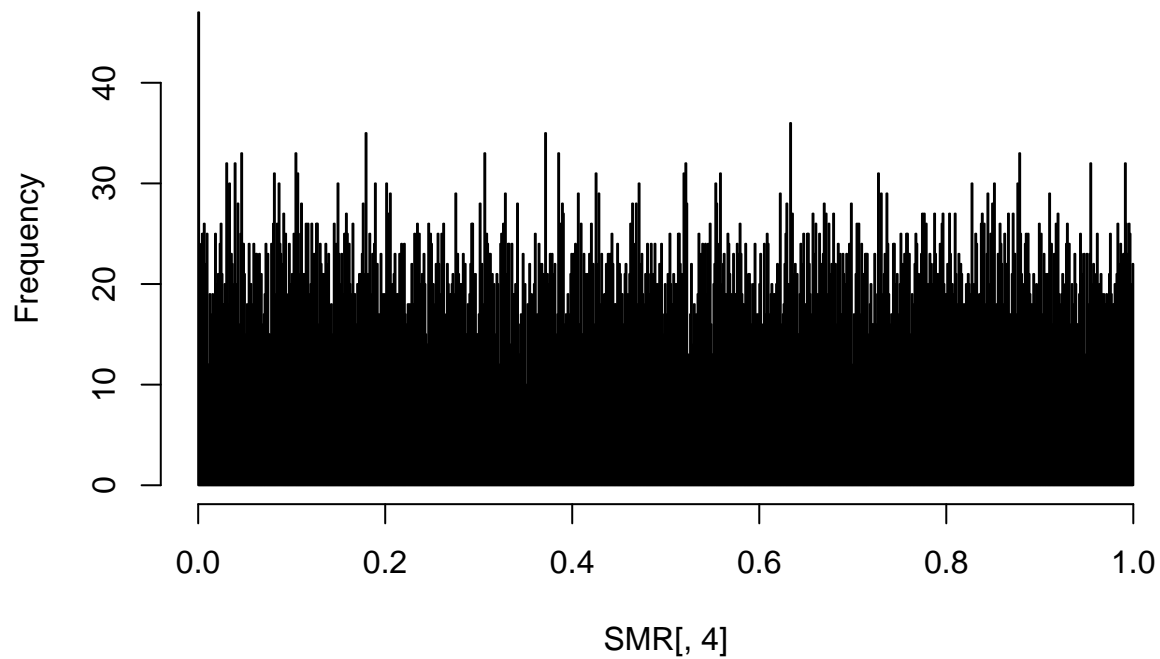
Histogram of p-values

Recall that under H0 p-values follow uniform distributions. If your data includes a fraction of alternative hypothesis you should see enrichment of low p-values.

Note that the fraction of Ha's is often small, so you need to use many bins in the histogram, in this case I specified 1,000,

```
hist(SMR[,4],1000,col=8)
```

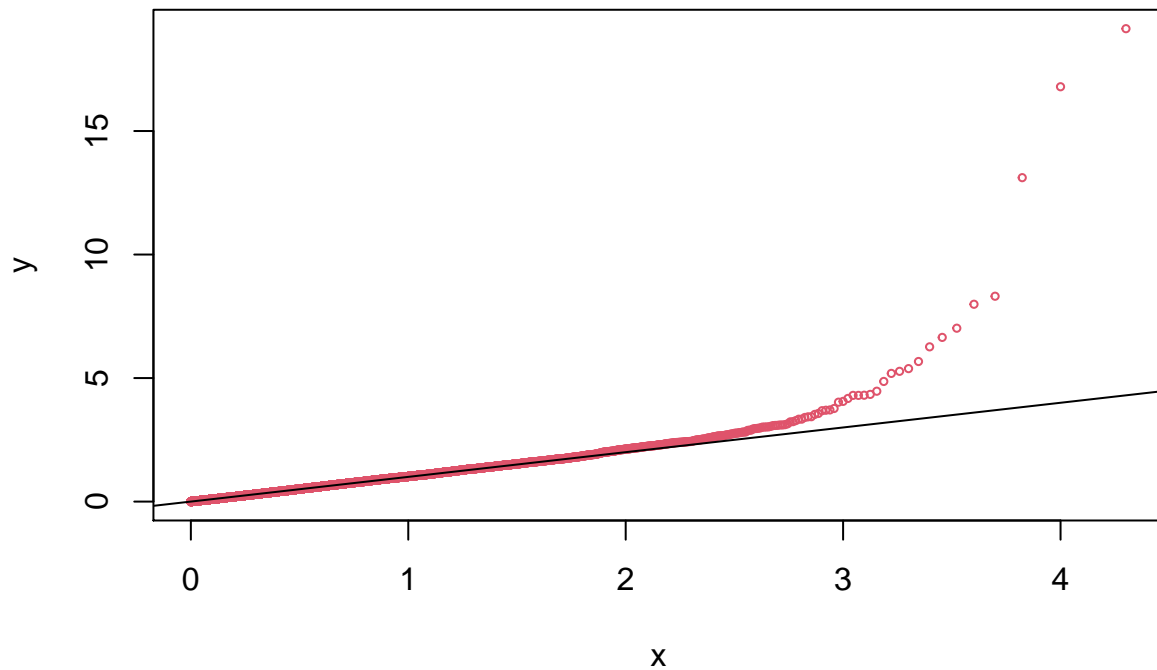# Histogram of SMR[, 4]

qqplot

Another, better way, to detect enrichment of low p-values is to compare the empirical quantiles of the p-values with theoretical quantiles for the uniform distribution.

```
pVal=sort(SMR[,4],decreasing=FALSE)
 expectedUnderH0=(1:length(pVal))/length(pVal)

 y= -log10(pVal)
 x= -log10(expectedUnderH0)

plot(y~x,cex=.5,col=2);abline(a=0,b=1)
```

```
## Using the qqman package
 #install.packages(pkg='qqman',repos='https://cran.r-project.org/')
 library(qqman)
```

```
##
## For example usage please run: vignette('qqman')
##
## Citation appreciated but not required:
## Turner, S.D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. biorXiv
##
```

```
 qq(SMR[,4])
```