Accumulation Test – Short Demo

Ang Li and Rina Foygel Barber

May 27, 2015

This code is a short demo of the accumulation test methods and code from the paper:

- Ang Li and Rina Foygel Barber, "Accumulation tests for FDR control in ordered hypothesis testing" (2015). Available from http://arxiv.org/abs/1505.07352 (http://arxiv.org/abs/1505.07352)

The demo was run using R version 3.2.0.

Setup

Load the functions for the accumulation test methods.

```
source('accumulation_test_functions.R')
```

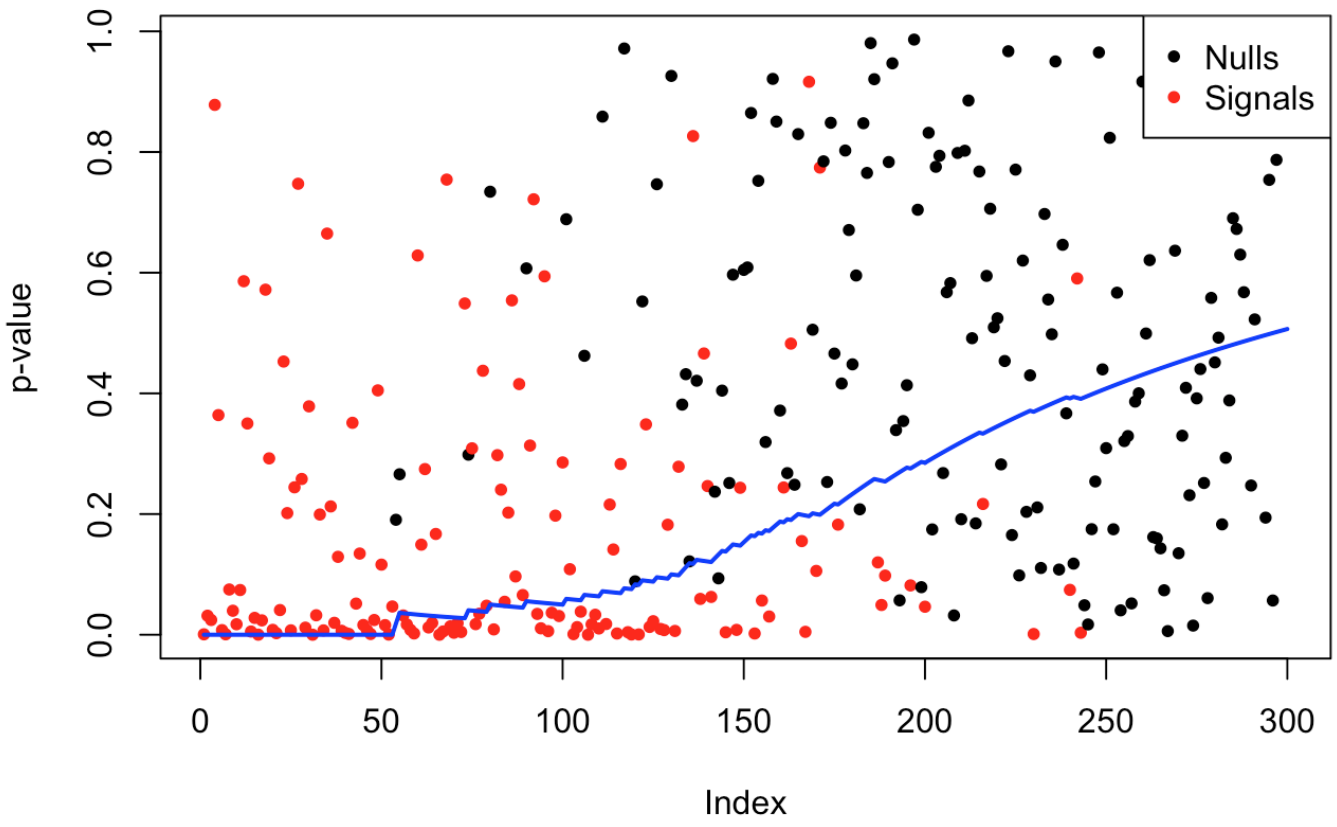
p-values

Generate an ordered list of p-values.

```
n=300
# probability of finding a true signal decays as we move along the sequence
TrueSignals=rbinom(n,1,1/(1+exp((1:n)/30-5)))

# now generate the p-values from Gaussian data
mu = 2; means = mu*TrueSignals # mu controls strength of true signals
zscores = rnorm(n) + means
pvals = 2*(1-pnorm(abs(zscores))) # 2-sided z-test
```

Plot the p-values. The curve shows the false discovery proportion as we move along the list.

```
plot(1:n, pvals, xlab = 'Index', ylab = 'p-value', col = 1+TrueSignals,pch=20)

legend('topright',legend=c('Nulls','Signals'),col=1:2,pch=20)

fdp=cumsum(1-TrueSignals)/(1:n)
points(1:n,fdp,type='l',col='blue',lwd=2)
```



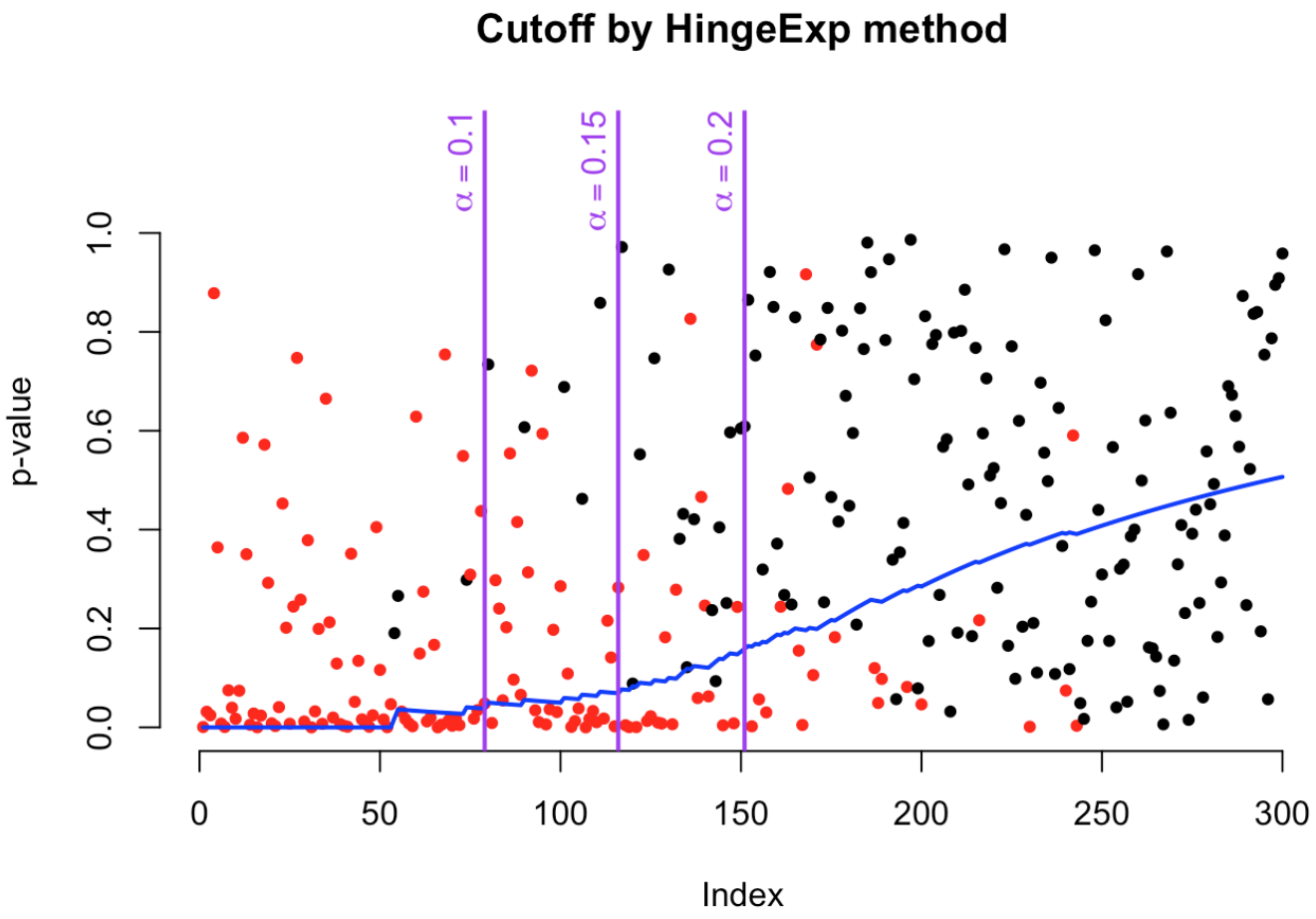
HingeExp method

For $\alpha = 0.1, 0.15, 0.2$, get the cutoff $\widehat{k}$ of accumulation test, using HingeExp method (with default parameter values).

```
alphas=c(0.1,0.15,0.2)
khats = HingeExp(pvals,alpha=alphas)
```

View results (code visible in .Rmd file):

```
## [1] Cutoff when alpha =  0.1 :  79
## [1] Cutoff when alpha =  0.15 :  116
## [1] Cutoff when alpha =  0.2 :  151
```



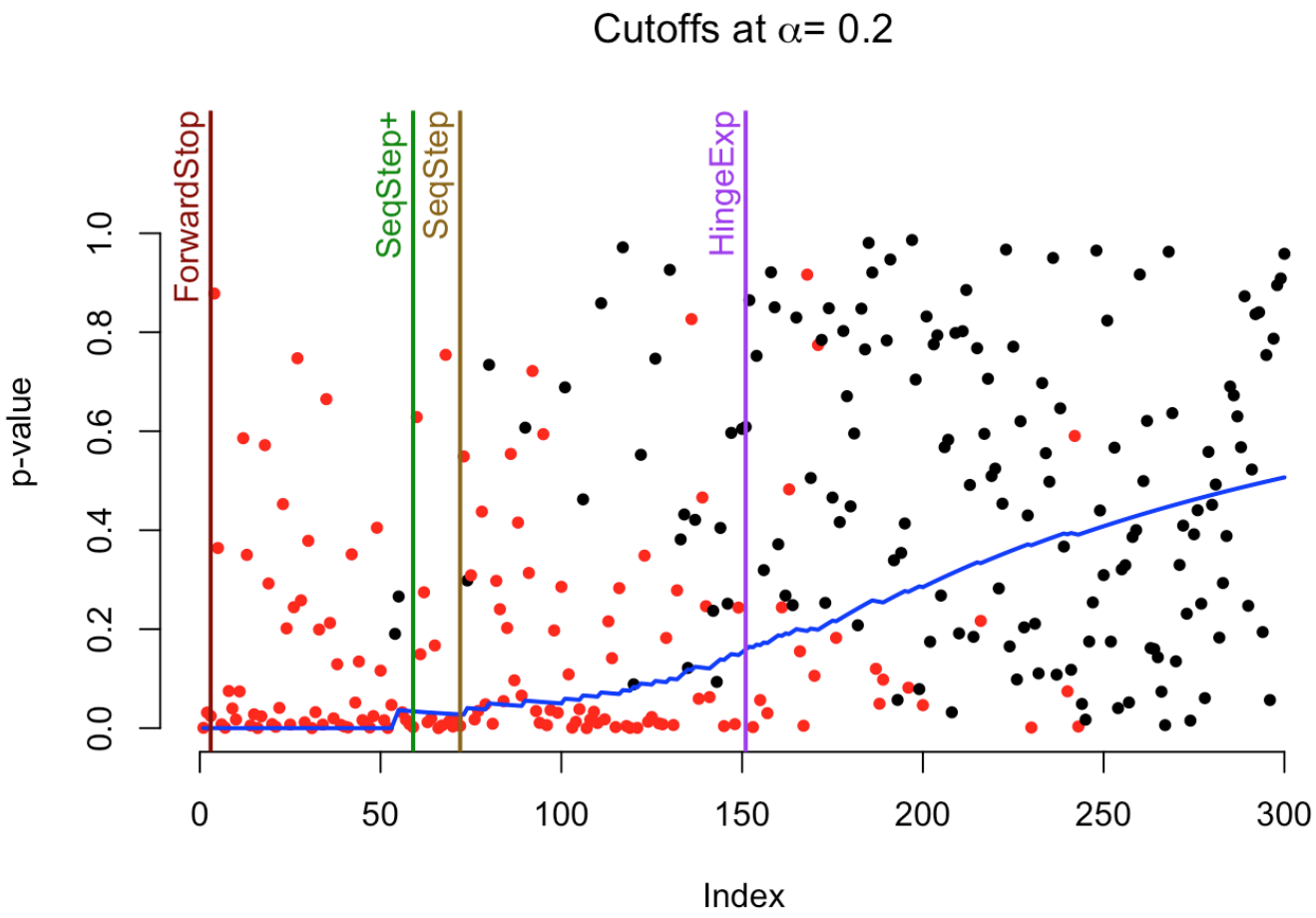
Comparing accumulation functions

For $\alpha = 0.2$, get the cutoff $\widehat{k}$ of accumulation tests, comparing the HingeExp method with the SeqStep and SeqStep+ (Barber and Candès 2014) methods and the ForwardStop (G'Sell et al. 2013) method (all with default parameter values).

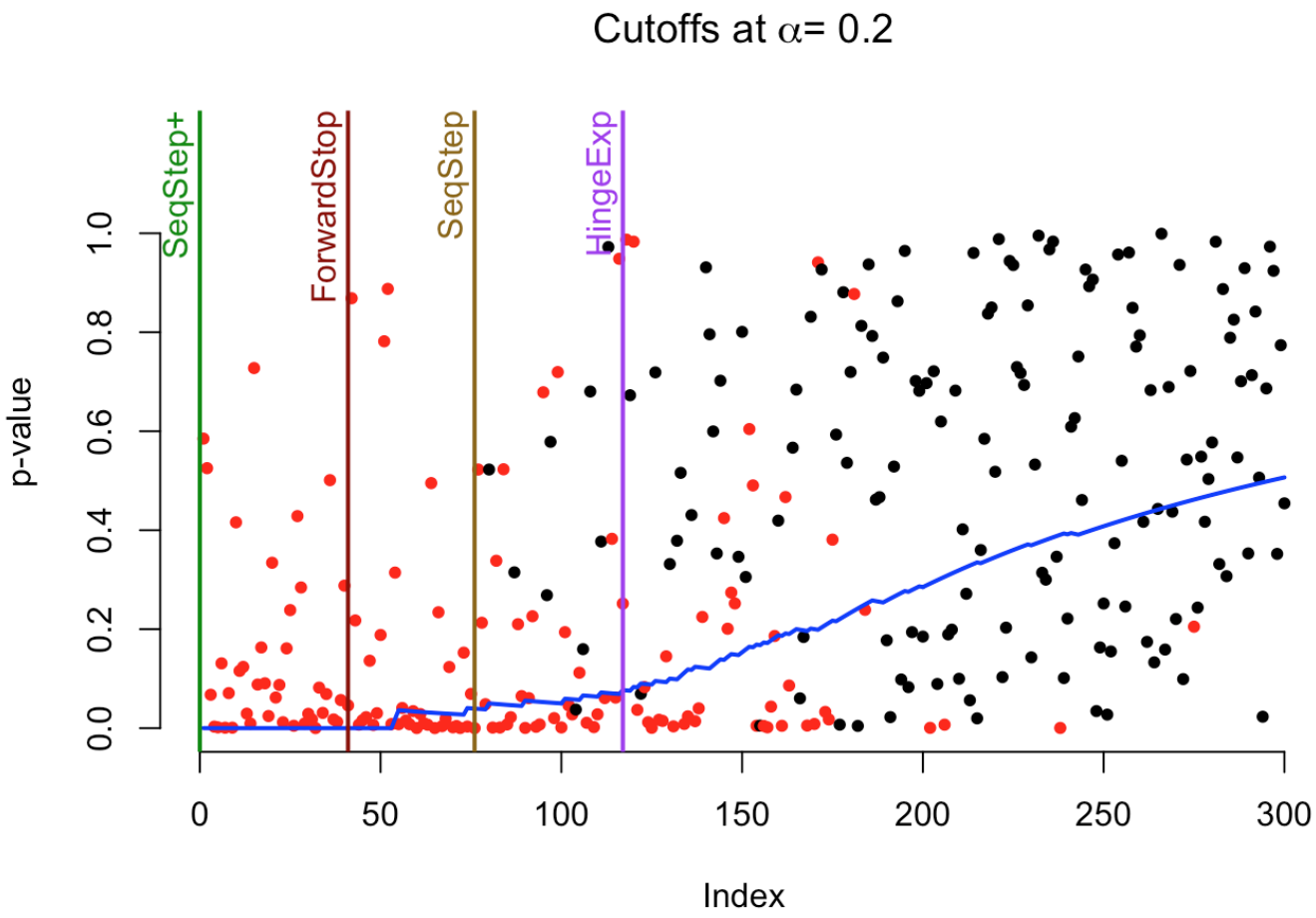
```
alpha = 0.2
khat = list()

khat$SeqStep = SeqStep(pvals,alpha)
khat$SeqStepPlus = SeqStepPlus(pvals,alpha)
khat$ForwardStop = ForwardStop(pvals,alpha)
khat$HingeExp = HingeExp(pvals,alpha)
```

View results (code visible in .Rmd file):

```
## [1] Cutoff when method is  SeqStep :  72
## [1] Cutoff when method is  SeqStep+ :  59
## [1] Cutoff when method is  ForwardStop :  3
## [1] Cutoff when method is  HingeExp :  151
```



The results will vary depending on the randomly generated p-values, and in some cases multiple methods will produce the same outcome. Below is a plot of the results when the same procedure is run again for a new random sequence of p-values generated in the same way (code visible in the .Rmd file).



Barber, Rina Foygel, and Emmanuel Candès. 2014. “Controlling the False Discovery Rate via Knockoffs.” ArXiv Preprint ArXiv:1404.5609.

G’Sell, Max Grazier, Stefan Wager, Alexandra Chouldechova, and Robert Tibshirani. 2013. “False Discovery Rate Control for Sequential Selection Procedures, with Application to the Lasso.” ArXiv Preprint ArXiv:1309.5352.