SCRIPT MOD6S4C: HIGHEST POSTERIOR DENSITY INTERVALS LABOR DATA APPLICATION

INSTRUCTOR: KLAUS MOELTNER

Load data

```
R> load("c:/Klaus/AAEC5126/R/data/normindepLabor.rda")
```

Define HPDI function

```
R> khpdi<-function(x,a,B){
 #
 # This generates a highest posterior density interval for draws in vector
 # "x".
 # The posterior of x, i.e. p(x|y), must be unimodal.
 # a is a number between 0 and 1,
 # and indicates the 100(1-a)% desired "level of confidence"
 # for the interval. Usually, a will be 0.05.
 # B is the number of bins you want to have for the underlying histogram
 # that captures empirical frequencies. The higher this number, the more
 # accurate, but slower, is the algorithm.
 int<-sort(x)
 l<-length(int) #total number of draws
 bw<-(max(int) - min(int))/(B-1)  #set binwidth s.t. we get B equally sized bins
 # one could increase that number for higher accuracy
 t<-seq(min(int),max(int),by=bw) #vector with bin thresholds
 h<-matrix(0,1,B) #will collect bin counts
 for (j in 1:(B-1)) {
 h[j]<-length(int[int >=t[j] & int<t[j+1]])
 }
 h[B]<-h[B]+1
 h<-h/l   #translate counts into frequencies -> MUST SUM TO 1!!
 #
 startmat<-0
 stopmat<-0

 for (i in 1:B) {
     int<-0
     j<-i
     while (j<B) {
         int<-int+h[j] #add up probabilities until we hit (1-a)
         if (int>(1-a)) {
             startx<-min(x)+(i-1)*bw
             stopx<-min(x)+j*bw
             startmat<-c(startmat, startx)
```

```
            stopmat<-c(stopmat, stopx)
            #j*bw indicates the lower bound of the jth bin,
            # this is a consevative choice for an upper bound
            break  #exit while loop
        }
        j<-j+1
    }
}
startmat<-startmat[2:length(startmat)]
stopmat<-stopmat[2:length(stopmat)]
#
# now we need to find the smallest of these intervals
rangemat<-stopmat-startmat
all<-cbind(rangemat, startmat, stopmat)
all <- all[order(all[,1]),] #sort smallest range to largest
#
low<-all[1,2] # lower bound for HDPI
up<-all[1,3] # upper x for HDPI
return(list(low,up))
}
```

HPDI FOR SINGLE PARAMETERS

```
R> R<-ncol(betamat)
R> b4<-betamat[4,]#number of children <6
R> a<-0.05 #level of significance (in a classical sense)
R> B<-500   #number of bins for HPDI evaluation
R> bounds4<-khpdi(b4,a,B)
R> low4<-bounds4[[1]]
R> up4<-bounds4[[2]]
R> #
R> b5<-betamat[5,]#number of children 6-18
R> a<-0.05 #level of significance (in a classical sense)
R> B<-500   #number of bins for HPDI evaluation
R> bounds5<-khpdi(b5,a,B)
R> low5<-bounds5[[1]]
R> up5<-bounds5[[2]]
R> #
R> b8<-betamat[8,]#number of children 6-18
R> a<-0.05 #level of significance (in a classical sense)
R> B<-500   #number of bins for HPDI evaluation
R> bounds8<-khpdi(b8,a,B)
R> low8<-bounds8[[1]]
R> up8<-bounds8[[2]]
```

The lower bound for the (95%) HPDI for the fractional effect of "number of children under the age of 6" on annual earnings" is -0.9192. The upper bound for the (95%) HPDI for the fractional effect of "number of children under the age of 6" on annual earnings" is -0.3366.

The lower bound for the (95%) HPDI for the fractional effect of "number of children age 6-18" on annual earnings" is -0.2075. The upper bound for the (95%) HPDI for the fractional effect of

"number of children age 6-18" on annual earnings" is -0.0209.

The lower bound for the (95%) HPDI for the fractional effect of "additional year of education" on annual earnings" is 0.0204. The upper bound for the (95%) HPDI for the fractional effect of "additional year of education" on annual earnings" is 0.1201.

Plot postrior with HPDI bounds

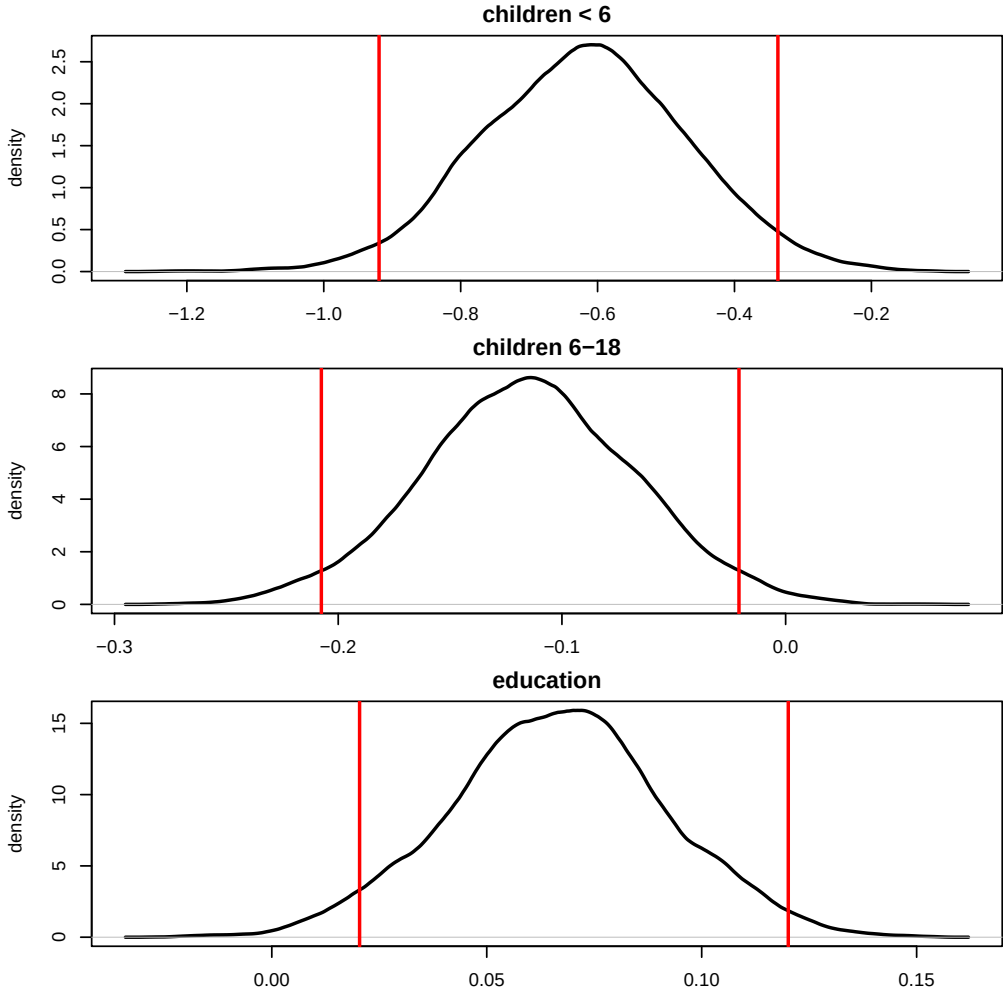

Figure 1. HPDIs for individual coefficients

Informal "Test" for restrictions involving several parameters

For an example, lets's examine if ten years of education can cancel out the effect of one small child.

```
R> # define R-matrix and c
R> Rm<-c(0, 0, 0, 1, 0, 0, 0, 10)
R> c<-0
R> #
R> #Get posterior density for Rb
R> Rb<-t(betamat) %*% as.matrix(Rm)
R> #
R> #Get HPDI
R> a<-0.05 #level of significance (in a classical sense)
R> B<-500   #number of bins for HPDI evaluation
R> bounds<-khpdi(Rb,a,B)
R> low<-bounds[[1]]
R> up<-bounds[[2]]
```

The lower bound for the (95%) HPDI for $10 * \beta 8 + \beta 4$ is -0.5058. The upper bound for the (95%) HPDI for $10 * \beta 8 + \beta 4$ is 0.591.

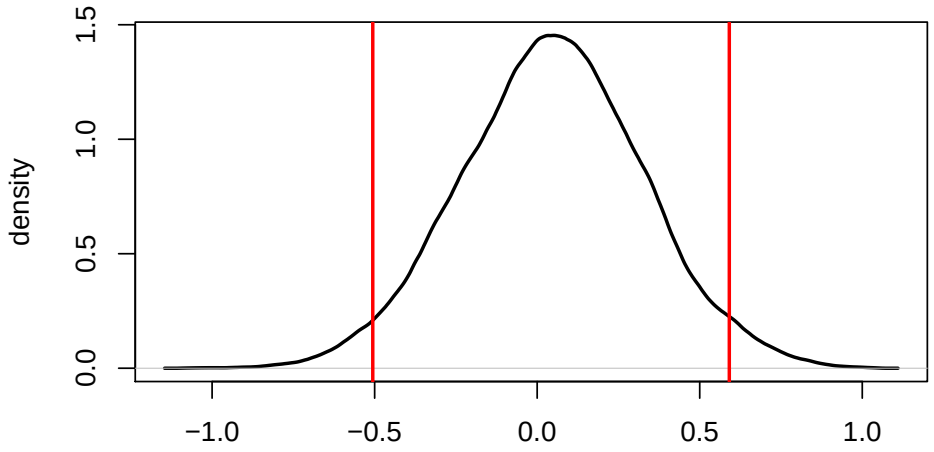

FIGURE 2. HPDI for $10 * \beta 8 + \beta 4$

```
R> proc.time()-tic
   user  system elapsed
   2.42    0.14    4.01
```