Instructions

(1) Homework must be typed and answered in the order given (problem 1(a)(b)(c)(d) first, problem 2(a)(b)... second, etc...)

(2) Undergrads and grads will answer all questions.

(3) Include in each part of the homework only the answer. R code and R output (without mistakes), must be included in the appendix to the question. For example, for question 1.a, write only the answer and your comments. The code and output for that part of the question will be in the appendix (the last part of question 1).

(4) No late homework under any circumstances.

(5) Write your name and ID this way: Last name, first name, UCLA ID, date, Homework number.

(6) Do not just give a number as an answer. For example, if asked for probability that posterior proportion is larger than 0.7, write $Prob(p > 0.7) = 0.3$, say and write comments or explanations if needed.

(7) The homework must be turned in in lecture (no mail box, no e-mail).

Problem 1. An experiment was performed to estimate the effect of beta-blockers on mortality of cancer patients. A group of patients were randomly assigned to treatment and control groups: out of 674 patients receiving the control, 39 died, and out of 680 receiving the treatment, 22 died. Assume that the outcomes are independent and binomially distributed, with probabilities of death of θ_1 and θ_2 under the control and treatment, respectively.

Let y_c be the number of patients that died in the control group and y_t the number of patients that died in the treatment group and let n_c be the number of patients in the control group and n_t the number of patients in the treatment group. Let also θ_1 be the population proportion of patients that died in the control group and θ_2 the population proportion of patients that died in the treatment group.

$n_c = 674$ $y_c = 39$ and $n_t = 680$ $y_t = 22$.

The likelihood function is

$$L \propto \theta_1^{39}(1-\theta_1)^{635}\theta_2^{22}(1-\theta_2)^{658}$$

Consider as prior a Dirichlet with $\alpha_1 = 1$ and $\alpha_2 = 1$. This means that $p(\theta_1, \theta_2) = 1$.

Under this non-informative prior, the posterior distribution is

$$p(\theta_1, \theta_2) \propto \theta_1^{39}(1-\theta_1)^{635}\theta_2^{22}(1-\theta_2)^{658}$$

(a) Obtain posterior simulations for the joint distribution of θ_1 and θ_2. Plot contour plots and a random sample from the joint distribution. **Hint:** For this and the next part, use as template, and modify as needed, the template used in Outline 7 (our class notes). Comment and interpret a little the plots.

The contour plot shows that highest posterior probability is around 0.06 for $\theta 1$ and 0.03 for $\theta 2$. Thus the contour plots say that the control group had a higher probability ($\theta_1 \approx 0.06$) of dying than the treatment group ($\theta_2 \approx 0.03$), suggesting that beta blockers seem to work for cancer patients. Notice that there seems to be more variability in the distribution of the $\theta 1$ than in that of $\theta 2$.

See figures 1 to 3

(b) The odds ratio is defined as $(\theta_2/(1-\theta_2))/(\theta_1/1-\theta_1))$. Summarize the posterior distribution for this estimand (mean, mode, median, standard deviation, 95% plot, etc..

Hint: to be able to obtain the posterior distribution of this function of θ_1 and θ_2, you will need first the marginal posterior distributions of θ_1 and θ_2 and then sample from them. Then, with the numbers obtained, you create the odds ratio and work with the distribution of this odds ratio.

I defined my odds ration as the inverse of the one given above, thus giving me the evidence of the control being between one and a half and two and a half times more likely to die than the treatment group. As you can see in

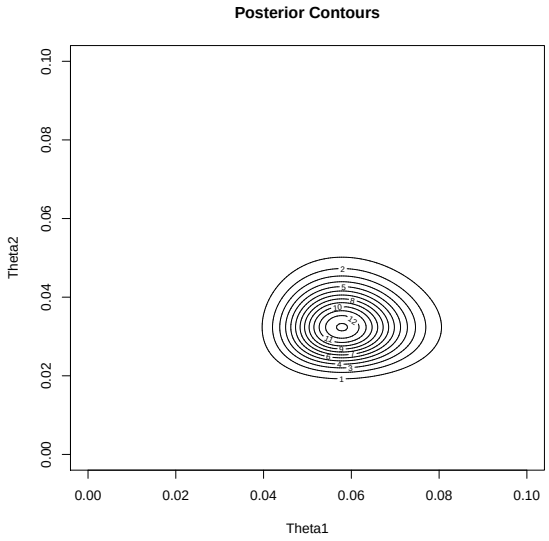

Figure 1: Contour plot of posterior distribution of theta1 and theta2, beta blockers

the histogram, (Figure 4) $(\theta_1/(1-\theta_1))/(\theta_2/1-\theta_2))$*, is overwhelmingly larger than 1: over 0.9 probability that the control group is more likely to die than the treatment group. If you did your ratio as defined in the question, most of your numbers will be less than 1... suggesting the same conclusion, that the treatment group is less than half times as likely to die as the control group.*

Thus, all the evidence points towards concluding that beta blockers work.

(c) Attach a well documented R code, i.e., add comments that illustrate what the code is doing, and separate your functions with headers that make them easy to find.

Appendix Question 1.

```
##########  Likelihood function ##############

likl.82 = function(a,b)      # a=theta1    b= theta2
{
y = c(39,22)              # yc=y[1]=39,  yt=y[2]=22
n= c(674,680)           #nc=n[1]=674, nt=n[2]=680
f=(aˆy[1])*( (1-a)ˆ(n[1]-y[1]) )*(bˆy[2])*( (1-b)ˆ(n[2]-y[2]) )
f
}

############ Posterior distribution ##############

post.82=function(a,b)
{
likl.82(a,b)
}

##########   Generate posterior contour plot #####
### Needs only function likl.82 and post.82 #######
```

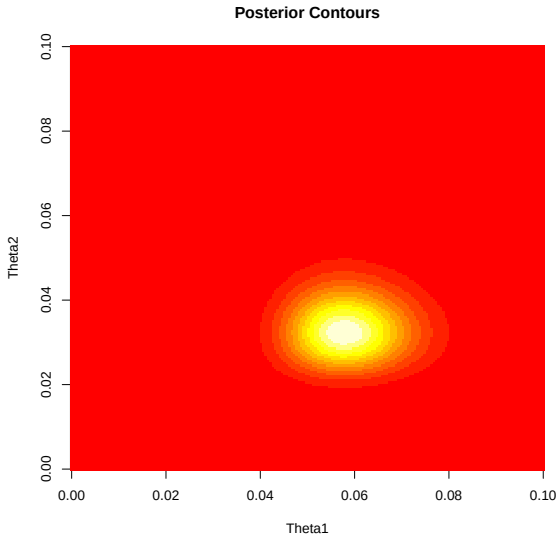

Figure 2: Image of contour plot of posterior distribution of theta1 and theta2, beta blockers

```
agrid= seq(0,0.1,length=200)
bgrid=seq(0,0.1,length=200)
p = matrix(0,nrow=200,ncol=200)  #initialize matrix for posterior

for(i in 1:200) {   # joint posterior distribution
   for(j in 1:200){
   p[i,j]=post.82(agrid[i],bgrid[j])
   }
      }

      ##### alternatively, exploit vector structure #####

# for(i in 1:200){
#    p[i,] = post.82(agrid[i],bgrid)
# }

  k=sum(p*0.00005)   # normalizing constant. You don't need for contour

 ############  Figures 1 and 2 are produced with these next 2 commands #######

  contour(agrid,bgrid,p/k,xlab="Theta1",ylab="Theta2",main="Posterior Contours")

#################################################################
#     Function init.rpost.82
 # requires: agrid, bgrid, p, p.agrid
 # sets up  marginal cdf for alpha:                cdfa
 #                 cond cdf for beta   | alpha:           cdfba
 # note: computing cdfba-inf for p(beta | alpha) would be too cumbersome,
```

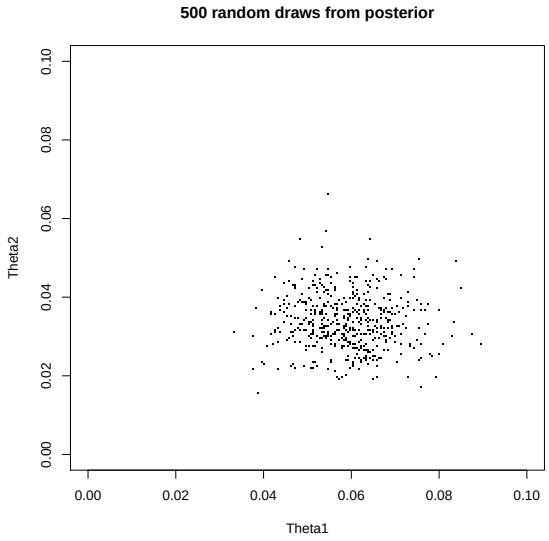

Figure 3: Random draws from the posterior distribution of theta1 and theta2, beta blockers

```
 # would need it on a whole grid of alphas !
 ##############################################################
 init.rpost.82 <- function()
 {
 da <- agrid[2]-agrid[1]              #step size on a grid
 k <- sum(p.agrid*da)        # probability in that interval
 cat("computing cdfa.inv....\n")
 cdfa <<- rep(0,200)   # initialize marginal cdf for alpha
 cdfa[1] <<- p.agrid[1]*da/k

 for ( i in 2:200)          #compute marginal cdf for alpha
   cdfa[i] <<- cdfa[i-1] + da*p.agrid[i]/k

      cdfba <<- matrix(0, nrow=200,ncol=200)   # initialize cdf for beta given alpha
         db <- bgrid[2] - bgrid[1]   #step size on bgrid
         k <- apply(p, 1, sum)*db     #vector of standardization constants
          cdfba[,1] <<- db*p[,1]/k
         for ( j in 2:200)
             cdfba[,j] <<- cdfba[, j-1]+db*p[,j]/k

           NULL
     }

p.agrid=apply(p,1,sum)
p.bgrid=apply(p,2,sum)

init.rpost.82()
```

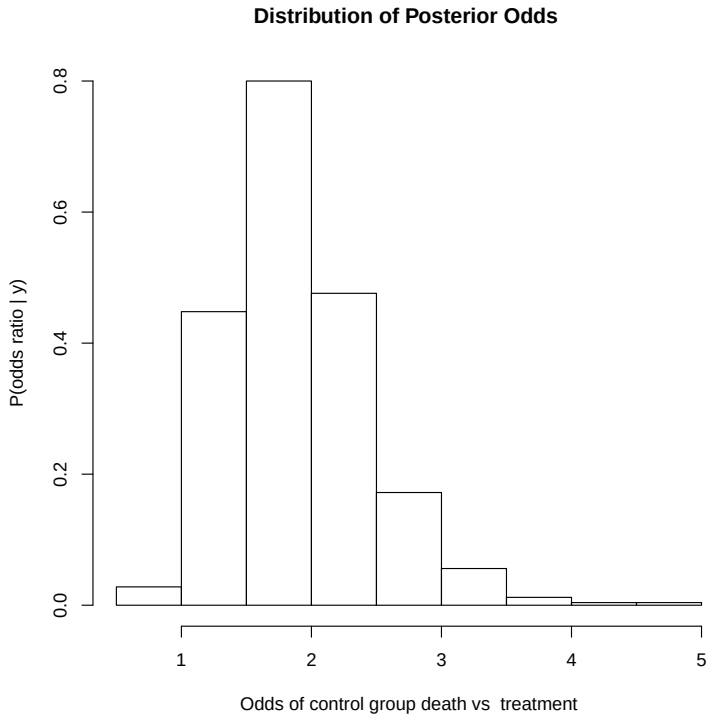

Figure 4: Distribution of posterior odds, beta blockers

```
################################################################
  #  Function rpost.82
  #
  #simulating a sample from the posterior using inverse cdf method
  # note:need to call init.rpost to setup cdfa & cdfba
  # see the Appendix to lesson 7 for an explanation of the inverse cdf method
  ################################################################
   rpost.82 <- function(n)
   {
    # Now generate p(a) and p(b |a)
   u1<-runif(n)
   u2<- runif(n)    #uniform r.v.
   a<- rep(0,n)   #initialization
   b<-rep(0,n)     #initialization
   for(j in 1:n){
       k <- (1:200)[cdfa>=u1[j]][1]  # k= first index st cdfa[k]>u1
       a[j] <-agrid[k]                          #a[j] =cdf-inv(u1[j])
       l<-(1:200)[cdfba[k,] >= u2[j]][1]     # l= first index st cdf[k]>u1
       b[j]<- bgrid[l]                          #b[j]=cdfab-inv(u2[j])
       }
       return(cbind(a,b))  # return a (n,2) matrix of a,b vectors.

     }
```

```
th=rpost.82(500)

plot(th[,1],th[,2],xlab="Theta1", ylab="Theta2",xlim=c(0,0.1),ylim=c(0,0.1), pch=".",main="500 random di

######## Odds Ratio #########

OR=function(th)
{
theta1=th[,1]
theta2=th[,2]
OR = (theta1/(1-theta1))/(theta2/(1-theta2))
OR
}
OR =OR(th)
hist(OR,xlab="Odds of control group dying more than treatment", ylab="P(odds ratio | y)",main="Distribu
```

Problem 2. This is a problem on material from Chapter 5 of Hoff's book. Thirty two students in a science class were randomly assigned to one of two study methods, A and B, so that $n_A = n_B = 16$ students were assigned to each method. After several weeks of study, students were examined on the course material with an exam designed to give an average score of 75 with a standard deviation of 10. The scores for the two groups are summarized by $\{\bar{y_A} = 75.2; s_A = 7.3\}$ and $\{\bar{y_B} = 77.5; s_B = 8.1\}$. Consider independent, conjugate normal prior distributions for each of θ_A and θ_B, with $\mu_0 = 75$ and $\sigma_0^2 = 100$ for both groups. For each $(\kappa_0, \upsilon_0) \in \{(1,1),(2,2),(4,4),(8,8),(16,16),(32,32)\}$, obtain $Pr(\theta_A < \theta_B \mid y_A, y_B)$ via Montecarlo sampling. Plot this probability as a function of $(\kappa_0 = \upsilon_0)$. Display how you may use this plot to convey the evidence that $\theta_A < \theta_b$ to people of a variety of prior opinions.

Hint: Do separately, for each group, the procedure described on page 77-78 of Hoffs book, to obtain random draws from the joint distribution and from there the marginal distribution of each theta. We have done draws like these before, in section 6.4 of our course notes, but applied to the other distributions. Read also section 6.4 to remind yourselves of what we did then.

Attach your R code.

Solution 1. Having $(\kappa_0 = \upsilon_0)$ with the same prior mean and prior variance, means that the belief is that both groups will perform the same. When $(\kappa_0 = \upsilon_0)$ = small, that belief is not very strong. When those numbers are very large, that belief is very strong.

We can see from the results obtained under lowest to highest $(\kappa_0 = \upsilon_0)$ (0.793 0.797 0.766 0.708 0.689 0.664) that the $Pr(\theta_A < \theta_B \mid y_A, y_B)$ decreases as the prior sample sizes increase, i.e. as the prior plays a bigger role in the posterior distribution. However, this probability is higher than 0.5 in all cases.

Thus, to people with a variety of prior opinions, I would say that no matter the strength of their prior beliefs that the two groups will perform the same (as conveyed by higher or lower $(\kappa_0 = \upsilon_0)$), the data would update those beliefs to a posterior that tells us that there is a relatively high probability that group A did worse than group B on average. The data leads to an update of their opinions pointing that the probability is almost 0.8 that group A did better than group B and is not lower than 0.5.

Appendix: R code

```
 ################# Prior ################
#prior for groupa A  and B
mu0=75
s20 =100

 k0.nu0 = matrix(c(1,1,2,2,4,4,8,8,16,16,32,32),ncol=2,byrow=T)
```

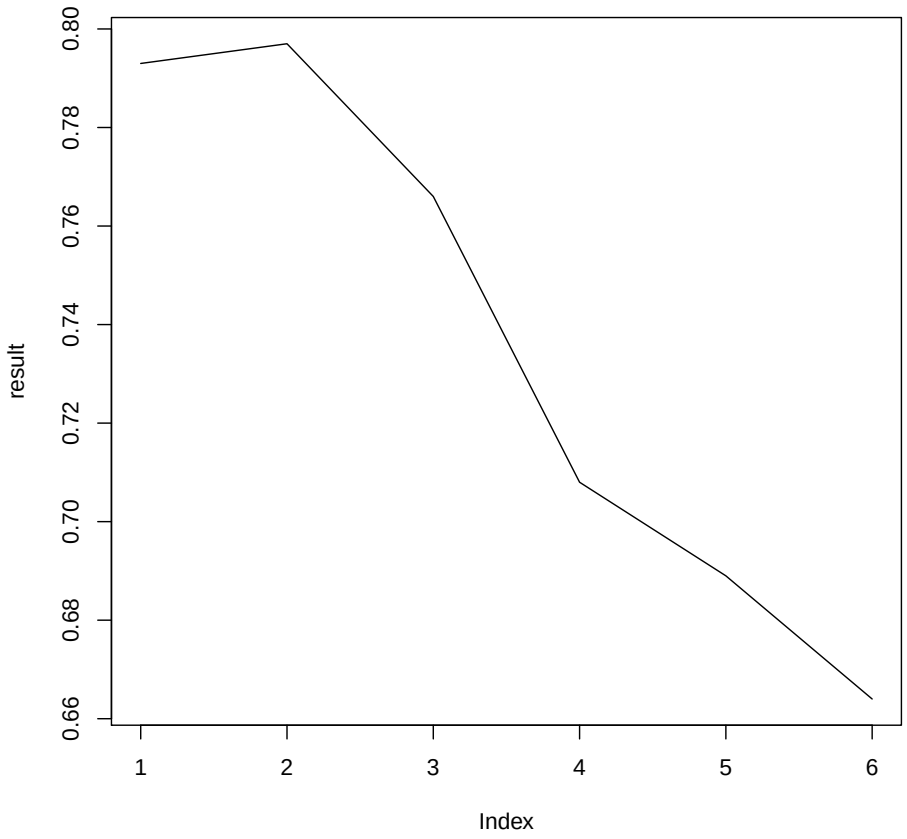

Figure 5: Sensitivity of posterior inference about superiority of group A to changes in prior beliefs

```
############## Data information ###########
# info from the data for group A
ybar.a=75.2;  s2.a = 7.3^2

n=16

# info from the data for group B
ybar.b=77.5; s2.b=8.1^2

########### Posterior mean and variance for each group ###########
## for each of the k0, nu0 and draws of 1000 random numbers #######
### from the posterior. Also compute the p(theta1 < theta 2) for each ####


result=rep(0,6)
for(i in 1:6){
```

```
    kn=k0.nu0[i,1]+n  ; nun=k0.nu0[i,2]+n
    mun.a=(k0.nu0[i,1]*mu0 + n*ybar.a)/kn
    mun.b=(k0.nu0[i,1]*mu0+n*ybar.b)/kn
    s2n.a=(k0.nu0[i,2]*s20 + (n-1)*s2.a + k0.nu0[i,1]*n*(ybar.a - mu0)^2 /(kn))/(nun)
    s2n.b=(k0.nu0[i,2]*s20 + (n-1)*s2.b + k0.nu0[i,1]*n*(ybar.b - mu0)^2 /(kn))/(nun)
s2.postsample.a = 1/rgamma(1000,nun/2,s2n.a*nun/2)
theta.postsample.a =rnorm(1000,mun.a, sqrt(s2.postsample.a/kn) )

     s2.postsample.b = 1/rgamma(1000,nun/2,s2n.b*nun/2)
theta.postsample.b =rnorm(1000,mun.b,sqrt(s2.postsample.b/kn))

result[i]=sum(theta.postsample.a < theta.postsample.b)/1000
}

 result
 0.793 0.797 0.766 0.708 0.689 0.664
```