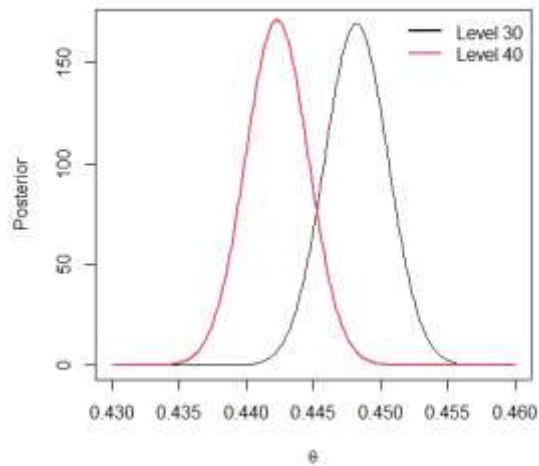


ST740 – Exam 1 – Solution

1a) The responses are the number of successes in a fixed number of trials so I selected likelihood $Y_j | \theta_j \sim \text{Binomial}(n_j, \theta_j)$ and conjugate prior $\theta_j \sim \text{Beta}(a, b)$, independent for $j=1, 2$. I set $a=b=1$ to give uninformative prior. From class, the posteriors are $\theta_j | Y_j \sim \text{Beta}(Y_j+a, n_j-Y_j+b)$.

1b) The plot below shows that the posterior mean retention rate is slightly higher for level 30, but there is overlap between the two posteriors.



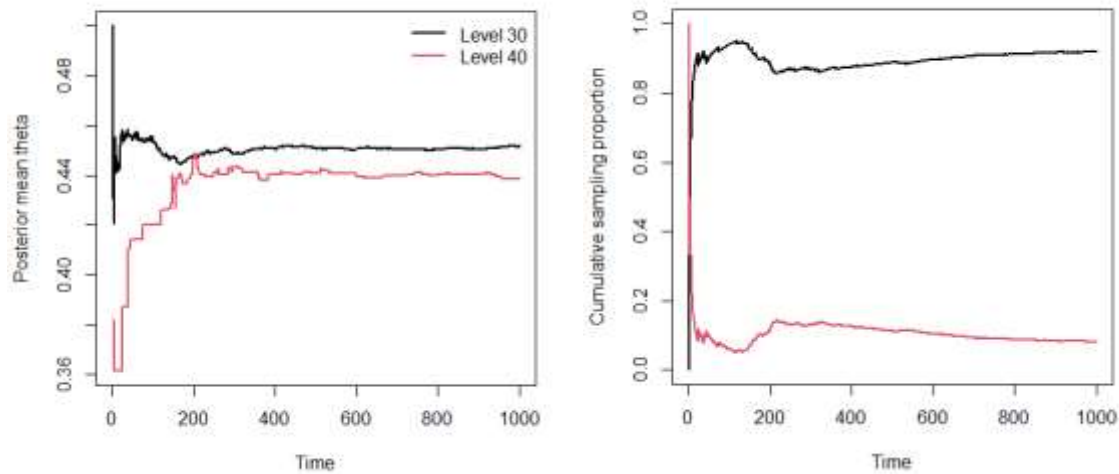
1c) I approximated the posterior probability that each level is optimal using 1M MC samples and found $P(\theta_1 > \theta_2 | Y_1, Y_2) = 0.96$ and $P(\theta_1 < \theta_2 | Y_1, Y_2) = 0.04$, so level 30 is likely optimal.

1d) I repeated 1c for several a with $a=b$ and found no sensitivity to the prior.

a	0.01	0.1	1	10	100
$P(\theta_1 > \theta_2 Y_1, Y_2)$	0.9629	0.9626	0.9628	0.9628	0.9625

```
2a) # Likelihood: Y_j ~ Binom(n_j, theta_j) j=1, 2
# Prior: theta_j ~ Beta(a, b)
# Inputs: Y1, Y2, n1, n2, a, b
# Output: group to be sampled, either 1 or 2
Thompson_sample <- function(Y1, n1, Y2, n2, a, b) {
  theta1 <- rbeta(1, Y1+a, n1-Y1+b) # Post sample from theta1
  theta2 <- rbeta(1, Y2+a, n2-Y2+b) # Post sample from theta2
  level <- ifelse(theta1 > theta2, 1, 2) # Level to be sampled
  return(level) }
```

2b) The plots below assume $a=b=1$ and plot the posteriors means, $(Y_j+a)/(n_j+a+b)$, and proportion of the cumulative samples assigned to each level, $n_j/(n_1+n_2)$, for $j=1,2$ by day. The posterior mean is consistently higher for level 30 than 40 so most users are assigned level 30.



2c) As shown in 2b, an advantage of Thompson sampling is that it explores both levels early when the optimal level is uncertain, but in the long-run most users are given the preferred level. Disadvantages are that implementing Thompson sampling brings some logistical challenges and it is possible that the algorithm gets stuck on a sub-optimal level.

```

# Code for 1b
n1 = 44700
Y1 = 20034
n2 = 45489
Y2 = 20119
a = b = 1

theta <- seq(0.43,0.46,0.0001)
plot(theta,dbeta(theta,Y1+a,n1-Y1+b),type="l",lwd=1,col=1,
      xlab=expression(theta),ylab="Posterior")
lines(theta,dbeta(theta,Y2+a,n2-Y2+b),lwd=2,col=2)
legend("topright",c("Level 30","Level 40"),lwd=2,col=1:2,bty="n")

# Code for 1c
theta1 <- rbeta(1000000,Y1+a,n1-Y1+b)
theta2 <- rbeta(1000000,Y2+a,n2-Y2+b)
mean(theta1>theta2)

# Code for 1d
ab <- c(0.01,0.1,1,10,100)
p1 <- ab
for(j in 1:length(ab)){
  theta1 <- rbeta(1000000,Y1+ab[j],n1-Y1+ab[j])
  theta2 <- rbeta(1000000,Y2+ab[j],n2-Y2+ab[j])
  p1[j] <-mean(theta1>theta2)
}
cbind(ab,p1)

# Code for 2a
# Likelihood:  $Y_j \sim \text{Binom}(n_j, \theta_j)$   $j=1,2$ 
# Prior:  $\theta_j \sim \text{Beta}(a,b)$ 
# Inputs: Y1,Y2,n1,n2,a,b
# Output: group to be sample, either 1 or 2
Thompson_sample <- function(Y1,n1,Y2,n2,a,b){
  theta1 <- rbeta(1,Y1+a,n1-Y1+b) # Posterior sample from theta1
  theta2 <- rbeta(1,Y2+a,n2-Y2+b) # Posterior sample from theta2
  level <- ifelse(theta1>theta2,1,2) # Level to be sampled
return(level)}

# Code for 2c
days <- 1000
p <- c(Y1/n1,Y2/n2)
n <- Y <- matrix(0,days+1,2)
for(t in 1:days){
  set.seed(919*t)
  lev <- Thompson_sample(Y[t,1],n[t,1],Y[t,2],n[t,2],a=1,b=1) # Pick level
  y <- rbinom(1,100,p[lev]) # Generate new data
  Y[t+1,] <- Y[t,] # Update current totals
  Y[t+1,lev] <- Y[t,lev] + y
  n[t+1,] <- n[t,]
  n[t+1,lev] <- n[t,lev] + 100
}
Y <- Y[-1,]; n <- n[-1,] # Remove time 0

post_mean <- (Y+a)/(n+a+b)
cum_n <- cbind(n[,1]/(n[,1]+n[,2]),n[,2]/(n[,1]+n[,2]))
matplot(post_mean,type="l",lwd=2,lty=1,col=1:2,
        xlab="Time",ylab="Posterior mean theta")
legend("topright",c("Level 30","Level 40"),lwd=2,col=1:2,bty="n")
matplot(cum_n,type="l",lwd=2,lty=1,col=1:2,
        xlab="Time",ylab="Cumulative sampling proportion")

```