lecture 2.2 example 4

preliminaries

```
# clear environment
rm(list = ls())

# set random seed
set.seed(123)
```

define params

```
#define grid size
nGridPoints = 50
pGrid = seq(from = 0, to = 1,length.out = nGridPoints)
gridSize = 1 / nGridPoints

# prior params
aPrior = 5
bPrior = 5
```

define simulation functions

```
computePost = function(data, prior) {
  nWater = sum(data)
  nData = length(data)
  likelihood = dbinom(x = nWater, size = nData, prob = pGrid)
  post = likelihood * prior
  post = post / ( sum(post) * gridSize )
  return(post)
}
```

load data

Assumed to be in the Desktop (in Mac system)

```
data = read.csv("~/Desktop/data_globe_tossing.csv")
```

prediction exercise

```
# compute posterior using first 50 observations
prior = dbeta(x = pGrid, shape1 = aPrior, shape2 = bPrior)
post = computePost(data$ball_1[1:50], prior)

# predict 1000 new observations
nPredict = 1000
theta = sample(pGrid, nPredict, replace = TRUE, prob = post)
```

```
prediction = rep(-1,nPredict)
for (i in 1:nPredict) {
  prediction[i] = rbinom(1, 1, theta[i])
}

# compare predicted vs actual
meanPredicted = mean(prediction)
actual = mean(data$ball_1[51:100])
paste ("mean water landings predicted =", meanPredicted)
```

```
## [1] "mean water landings predicted = 0.487"
```

```
paste ("actual out of sample water landings =", actual)
```

```
## [1] "actual out of sample water landings = 0.46"
```