```
# Suppose we are giving two students a multiple-choice exam with 40 questions,
# where each question has four choices. We don't know how much the students
# have studied for this exam, but we think that they will do better than just
# guessing randomly.
# 1) What are the parameters of interest?
# 2) What is our likelihood?
# 3) What prior should we use?
# 4) What is the prior probability P(theta>.25)? P(theta>.5)? P(theta>.8)?
# 5) Suppose the first student gets 33 questions right. What is the posterior
#    distribution for theta1? P(theta1>.25)? P(theta1>.5)? P(theta1>.8)?
#    What is a 95% posterior credible interval for theta1?
# 6) Suppose the second student gets 24 questions right. What is the posterior
#    distribution for theta2? P(theta2>.25)? P(theta2>.5)? P(theta2>.8)?
#    What is a 95% posterior credible interval for theta2?
# 7) What is the posterior probability that theta1>theta2, i.e., that the
#    first student has a better chance of getting a question right than
#    the second student?

############
# Solutions:

# 1) Parameters of interest are theta1=true probability the first student
#    will answer a question correctly, and theta2=true probability the second
#    student will answer a question correctly.

# 2) Likelihood is Binomial(40, theta), if we assume that each question is
#    independent and that the probability a student gets each question right
#    is the same for all questions for that student.

# 3) The conjugate prior is a beta prior. Plot the density with dbeta.
theta=seq(from=0,to=1,by=.01)
plot(theta,dbeta(theta,1,1),type="l")
plot(theta,dbeta(theta,4,2),type="l")
plot(theta,dbeta(theta,8,4),type="l")

# 4) Find probabilities using the pbeta function.
1-pbeta(.25,8,4)
1-pbeta(.5,8,4)
1-pbeta(.8,8,4)

# 5) Posterior is Beta(8+33,4+40-33) = Beta(41,11)
41/(41+11)  # posterior mean
33/40       # MLE

lines(theta,dbeta(theta,41,11))

# plot posterior first to get the right scale on the y-axis
plot(theta,dbeta(theta,41,11),type="l")
lines(theta,dbeta(theta,8,4),lty=2)
# plot likelihood
lines(theta,dbinom(33,size=40,p=theta),lty=3)
# plot scaled likelihood
lines(theta,44*dbinom(33,size=40,p=theta),lty=3)

# posterior probabilities
1-pbeta(.25,41,11)
1-pbeta(.5,41,11)
1-pbeta(.8,41,11)

# equal-tailed 95% credible interval
qbeta(.025,41,11)
qbeta(.975,41,11)

# 6) Posterior is Beta(8+24,4+40-24) = Beta(32,20)
```

```
32/(32+20)  # posterior mean
24/40       # MLE

plot(theta,dbeta(theta,32,20),type="l")
lines(theta,dbeta(theta,8,4),lty=2)
lines(theta,44*dbinom(24,size=40,p=theta),lty=3)

1-pbeta(.25,32,20)
1-pbeta(.5,32,20)
1-pbeta(.8,32,20)

qbeta(.025,32,20)
qbeta(.975,32,20)

# 7) Estimate by simulation: draw 1,000 samples from each and see how often
#    we observe theta1>theta2

theta1=rbeta(1000,41,11)
theta2=rbeta(1000,32,20)
mean(theta1>theta2)


# Note for other distributions:
# dgamma,pgamma,qgamma,rgamma
# dnorm,pnorm,qnorm,rnorm
```