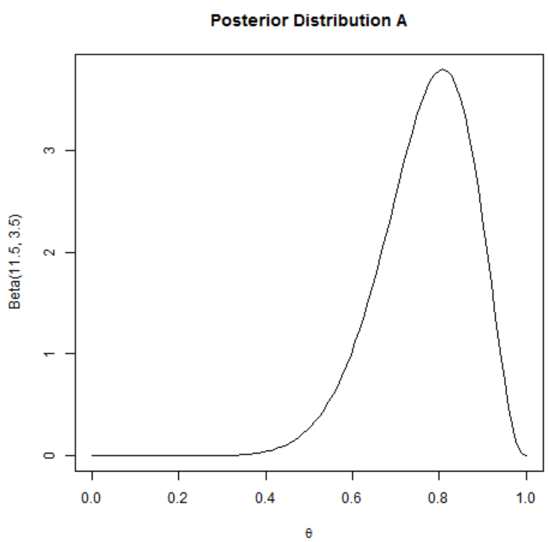
STA 360: Lab 1

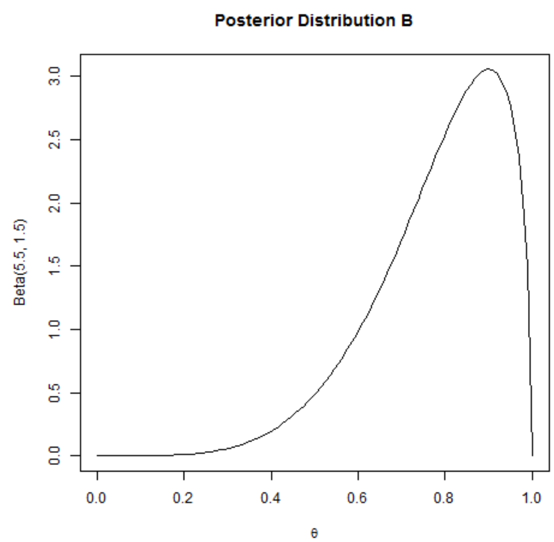
Michael Lin

January 18, 2015

1. The following is the derivation for the general form of the posterior distributions. For clarity, we used θ instead of p as the probability of success. The plots of the posterior distributions are included below as well.

$$\begin{aligned}
p(\theta|x_{1:n}) &\propto p(\theta)p(x_{1:n}|\theta) \\
&= \theta^{a-1}(1-\theta)^{b-1} \cdot \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \\
&= \theta^{a+\sum x_i - 1}(1-\theta)^{b+n-\sum x_i - 1} \\
&\propto \mathrm{Beta}(\theta|a+\sum_{i=1}^{n} x_i, b+n-\sum_{i=1}^{n} x_i)
\end{aligned}$$





2. The probabilities that the processes are successful at least 80 % of the time are as follows:

$$\mathbb{P}(\theta_A \geq 0.8) = 1 - \texttt{pbeta}(0.8, 11.5, 3.5) = 0.4204$$

$$\mathbb{P}(\theta_B \geq 0.8) = 1 - \texttt{pbeta}(0.8, 5.5, 1.5) = 0.5355$$

3. By taking large samples from posterior distributions of θ_A and θ_B, the probability that solution B truly has a higher success rate than solution A is about 0.57. The exact integral to compute this probability is:

$$\mathbb{P}(\theta_A < \theta_B|x_{A,1:n}, x_{B,1:n}) = \int_0^1 \int_0^{\theta_B} p(\theta_A, \theta_B|x_{A,1:n}, x_{B,1:n}) d\theta_A d\theta_B$$

4. Based on the fact that the probability that solution B truly has a higher success rate than solution A is about 0.57, she should use solution B.

See below for R code written to perform above analyses and calculations.

```
## Part 1
a.prior = 1/2;
b.prior = 1/2;
A.success = 11;
A.total = 14;
B.success = 5;
B.total = 6;

x.A = seq(0, 1, length=100);
hx.A = dbeta(x.A, a.prior+A.success, b.prior+A.total-A.success);

x.B = seq(0, 1, length=100);
hx.B = dbeta(x.B, a.prior+B.success, b.prior+B.total-B.success);

png("densityA.png");
plot(x.A, hx.A, type="l", lty=1, xlab=expression(theta),
     ylab="Beta(11.5, 3.5)", main="Posterior Distribution A");
dev.off()

png("densityB.png");
plot(x.B, hx.B, type="l", lty=1, xlab=expression(theta),
     ylab="Beta(5.5, 1.5)", main="Posterior Distribution B");
dev.off()


## Part 2
prob.A = 1-pbeta(0.8, a.prior+A.success, b.prior+A.total-A.success);
prob.B = 1-pbeta(0.8, a.prior+B.success, b.prior+B.total-B.success);


## Part 3
numsample = 1000000
sample.A = rbeta(numsample, a.prior+A.success, b.prior+A.total-A.success);
sample.B = rbeta(numsample, a.prior+B.success, b.prior+B.total-B.success);
counter = 0
for(i in 1:numsample) {
  if (sample.A[i] < sample.B[i]) counter = counter + 1;
}

blah = counter/numsample;
```