

```
library(latex2exp)
set.seed(0)
```

#1

a.

In deciding Monte Carlo sample size, to make sure it is sufficient so that the answers are correct to within three decimal places with 95% probability, the following must be true:

$$2\sqrt{\sigma^2/S} \leq 0.001$$

```
# store data and parameters
y_a <- c(12, 9, 12, 14, 13, 13, 15, 8, 15, 6)
y_b <- c(11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7)
a_a <- 120
b_a <- 10
a_b <- 12
b_b <- 1
n_a <- length(y_a)
sy_a <- sum(y_a)
n_b <- length(y_b)
sy_b <- sum(y_b)
```

```
smc <- 100000000 # 10 million
theta_a_mc <- rgamma(smc, a_a + sy_a, b_a + n_a)
y_a_predmc <- rpois(smc, theta_a_mc)

theta_b_mc <- rgamma(smc, a_b + sy_b, b_b + n_b)
y_b_predmc <- rpois(smc, theta_b_mc)

p1 <- mean(theta_a_mc > theta_b_mc)
p2 <- mean(y_a_predmc > y_b_predmc)

print(paste0('The Monte Carlo sample size is ', smc))
```

```
## [1] "The Monte Carlo sample size is 1e+07"
```

```
print(paste0('The probability that theta_A > theta_B given y_A and y_B is ', format(round(p1, 3))))
```

```
## [1] "The probability that theta_A > theta_B given y_A and y_B is 0.995"
```

```
print(paste0('The probability that predicted Y_A > predicted Y_A given y_A and y_B is ',
format(round(p2, 3))))
```

```
## [1] "The probability that predicted Y_A > predicted Y_A given y_A and y_B is 0.698"
```

```
# Monte Carlo sample size
var <- var(theta_a_mc > theta_b_mc)
min_sample_size <- var / ((0.001/2)^2)
print(min_sample_size)
```

```
## [1] 18342.31
```

Therefore, the selection of 10 million as the Monte Carlo sample size was sufficient to make sure that the answers are correct to within three decimal places with 95% probability.

b.

```
prob_theta <- list()
prob_pred <- list()
n_0_seq <- seq(1:10)

for (n in n_0_seq){
  theta_a_mc <- rgamma(smc, a_a + sy_a, b_a + n_a)
  y_a_predmc <- rpois(smc, theta_a_mc)

  theta_b_mc <- rgamma(smc, a_b*n + sy_b, n + n_b)
  y_b_predmc <- rpois(smc, theta_b_mc)

  p1 <- mean(theta_a_mc > theta_b_mc)
  p2 <- mean(y_a_predmc > y_b_predmc)

  prob_theta <- append(prob_theta, p1)
  prob_pred <- append(prob_pred, p2)
}
```

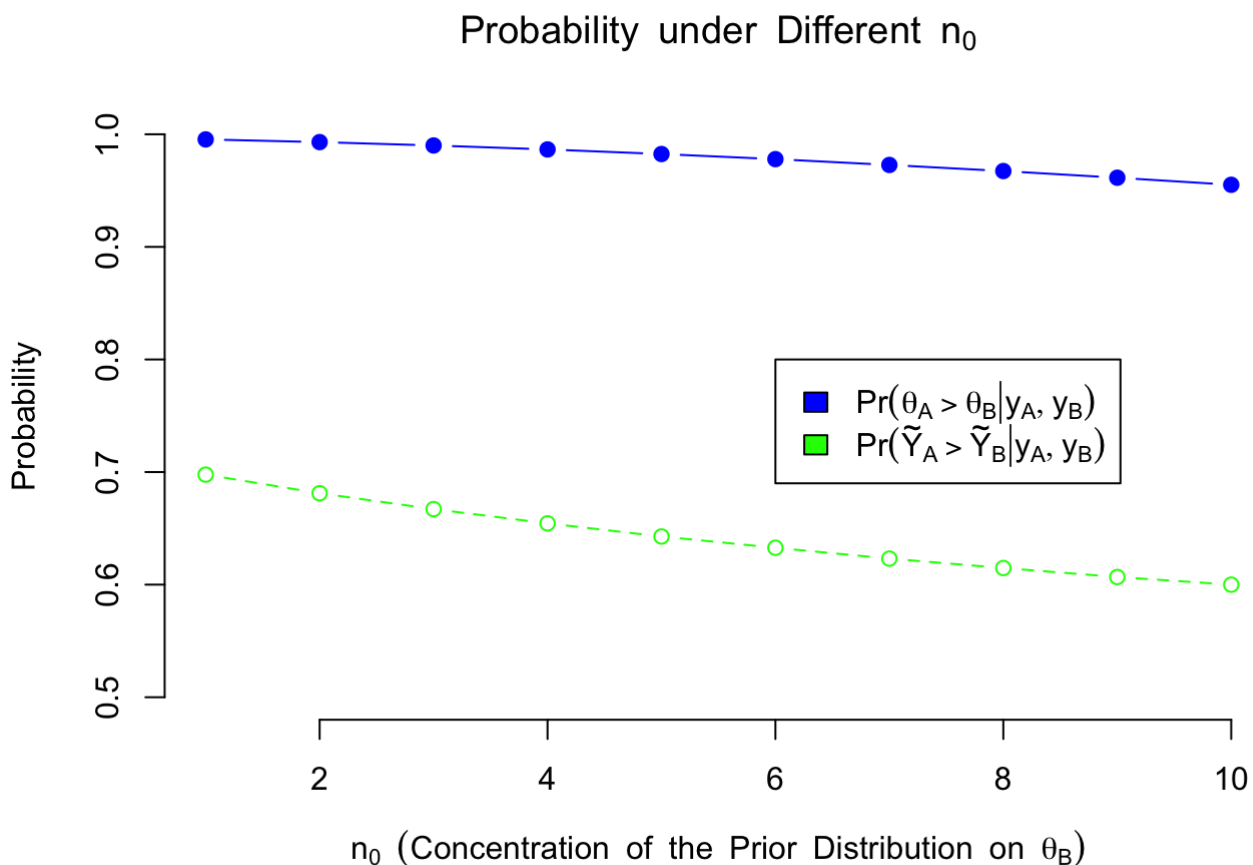
```

plot(n_0_seq, prob_theta,
     col = 'blue',
     type = "b",
     frame = FALSE,
     pch = 19,
     main = TeX('$Probability \\ under \\ Different \\ n_0$'),
     xlab = TeX('$n_0 \\ (Concentration \\ of \\ the \\ Prior \\ Distribution \\ on \\ \\theta_B)$'),
     ylab = TeX('$Probability$'),
     ylim = c(0.5, 1))

points(n_0_seq, prob_pred,
       col = 'green',
       type = "b",
       lty = 2)

legend(6, 0.8,
      legend = c(TeX('$Pr(\\theta_A > \\theta_B | y_A, y_B)$'), TeX('$Pr(\\tilde{Y}_A > \\tilde{Y}_B | y_A, y_B)$')),
      fill = c('blue', 'green'))

```



```
print(prob_theta[1])
```

```
## [[1]]  
## [1] 0.9954453
```

```
print(prob_theta[10])
```

```
## [[1]]  
## [1] 0.9551444
```

```
print(prob_pred[1])
```

```
## [[1]]  
## [1] 0.6977145
```

```
print(prob_pred[10])
```

```
## [[1]]  
## [1] 0.599989
```

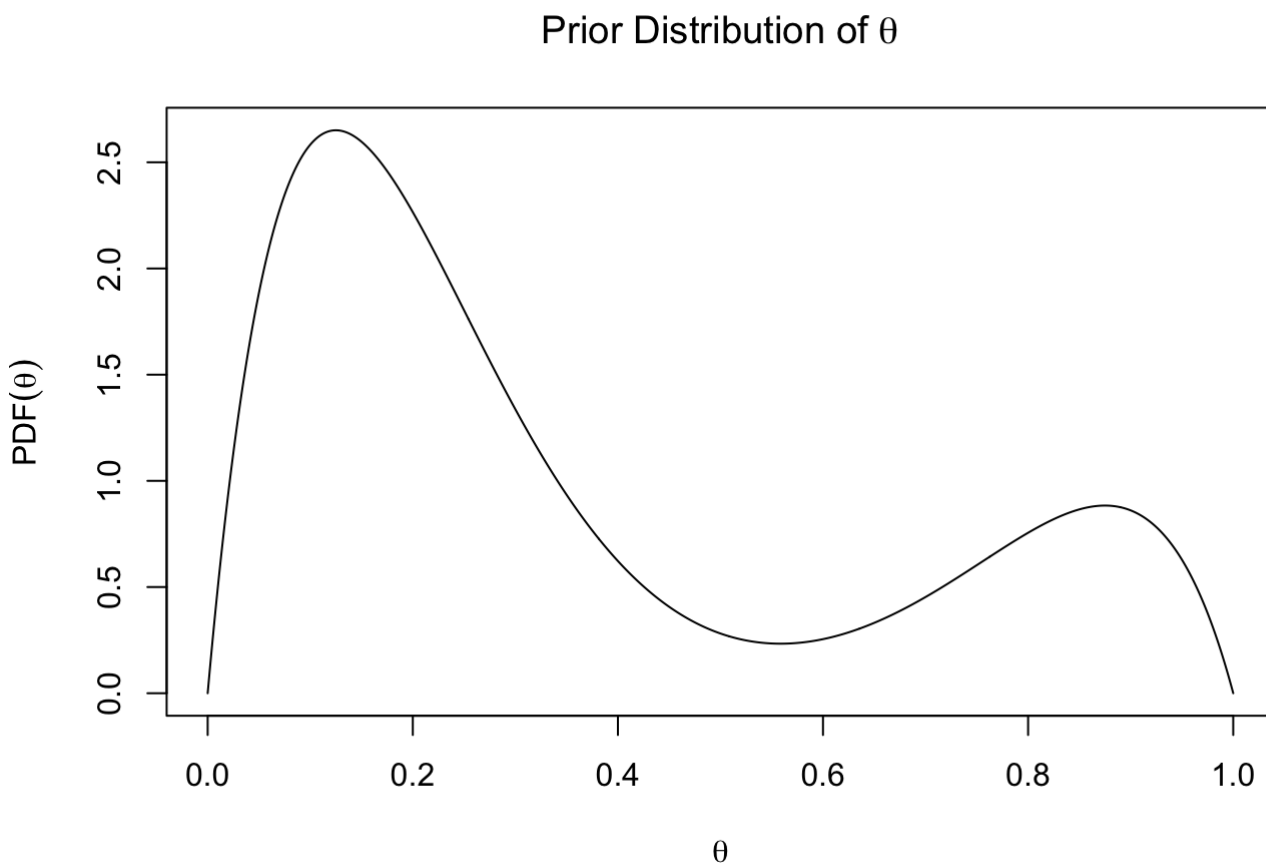
We know that as n_0 increases, the concentration of the prior distribution $\theta_B \sim \text{gamma}(12 \times n_0, n_0)$ would increase (higher peak and lower variance). According to the graph, we see that both $\Pr(\theta_A > \theta_B | y_A, y_B)$ and $\Pr(\tilde{Y}_A > \tilde{Y}_B | y_A, y_B)$ decreases when n_0 increases from 1 to 10. However, $\Pr(\tilde{Y}_A > \tilde{Y}_B | y_A, y_B)$ decreases by more (from 0.698 to 0.600) than $\Pr(\theta_A > \theta_B | y_A, y_B)$ (from 0.995 to 0.955). Therefore, $\Pr(\tilde{Y}_A > \tilde{Y}_B | y_A, y_B)$ is more sensitive to the concentration of the prior distribution than $\Pr(\theta_A > \theta_B | y_A, y_B)$.

#2

a.

```
# plot prior theta distribution
thetas <- seq(0, 1, length.out = 1000)
p_theta <- (1/4) * (gamma(10) / (gamma(2) * gamma(8))) * (3 * thetas * (1 - thetas)^7 +
  thetas^7 * (1-thetas))

plot(thetas, p_theta,
     type = "l",
     main = TeX("Prior Distribution of \\theta"),
     xlab = TeX("\\theta"),
     ylab = TeX('$PDF(\\theta)$'))
```



As we can observe from the graph, the prior distribution of θ is bimodal with peaks at around $\theta = 0.1$ and $\theta = 0.9$. This makes sense because the scientists have the prior knowledge that an experimental machine in a lab is either fine, in which case it would have a low failure rate (small θ), or comes from a bad batch of machines, in which case it would have a high failure rate (large θ). Therefore, their beliefs center around a comparatively small θ at around 0.1, and a comparatively large θ at around 0.9.

```
# find mode
thetas[which.max(p_theta)]
```

```
## [1] 0.1251251
```

The scientists think it is more likely that their machine is fine, which is revealed by the fact that the mode (highest peak) of the θ distribution is at 0.125, corresponding to the belief that the experimental machine in a lab is fine, in which case it would have a low failure rate (small θ).

C.

```
# plot posterior theta distribution
n <- 4
y_i <- 0:4
colors <- c('red', 'orange', 'yellow', 'green', 'cyan')

for (y in y_i){
  post_p_theta <- thetas^(y + 1) * (1 - thetas)^(7 + n - y) / beta(2 + y, 8 + n - y) +
  thetas^(y + 7) * (1 - thetas)^(1 + n - y) / (beta(8 + y, 2 + n - y))

  if (y == 0){
    plot(thetas, post_p_theta,
         type = "l",
         col = colors[1],
         main = TeX("Posterior Distribution of \\theta"),
         xlab = TeX("\\theta"),
         ylab = TeX('$PDF(\\theta)$'),
         ylim = c(0, 5))
  }
  else{
    lines(thetas, post_p_theta, col = colors[y])
  }
}

legend(0.4, 5,
      legend = c(TeX('$\\Sigma y_i = 0$'),
                  TeX('$\\Sigma y_i = 1$'),
                  TeX('$\\Sigma y_i = 2$'),
                  TeX('$\\Sigma y_i = 3$'),
                  TeX('$\\Sigma y_i = 4$')),
      fill = colors,
      cex = .8)
```

Posterior Distribution of θ 