STA 602. HW07

Yicheng Shen

10/21/2022

5.1

a. We know that

$$
\begin{aligned}
\theta|\sigma^2, y_1 \dots y_n &\sim N(\mu_n, \sigma^2/\kappa_n) \\
1/\sigma^2|y_1 \dots y_n &\sim Gamma(\nu_n/2, \nu_n\sigma_n^2/2) \\
\mu_n &= \frac{\kappa_0}{\kappa_0+n}\mu_0 + \frac{n}{\kappa_0+n}\bar{y} = \frac{\kappa_0}{\kappa_n}\mu_0 + \frac{n}{\kappa_n}\bar{y} \\
\nu_n &= \nu_0 + n \\
\sigma_n^2 &= 1/\nu_n + [\nu_0\sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_n}(\bar{y}-\mu_0)^2]
\end{aligned}
$$

```
school1 <- read.table("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/school1.dat")$V1
school2 <- read.table("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/school2.dat")$V1
school3 <- read.table("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/school3.dat")$V1
```

The following function computes the posterior distributions and obtains 5000 MC draws in order to get posterior mean and 95% CI.

```
# prior guess
mu0 <- 5; sigma02 <- 4
k0 <- 1; nu0 <- 2

compute_post <- function(school_data, S)
{
  n = length(school_data)
  ybar = mean(school_data)
  s2 = var(school_data)

  kn = k0 + n
  nun = nu0 + n
  mun = (k0 * mu0 + sum(school_data)) / kn
  sigma_n2 = (1 / nun) * (nu0 * sigma02 + (n - 1) * s2 + ((k0 * n) / kn) * (ybar - mu0)^2)

  sigma_2_draw <- 1 / rgamma(S, nun / 2, nun * sigma_n2 / 2)
  theta_draw <- rnorm(S, mun, sqrt(sigma_2_draw / kn))

  sigma_mean <- mean(sqrt(sigma_2_draw))
  theta_mean <- mean(theta_draw)

  sigma_quantile <- quantile(sqrt(sigma_2_draw), probs = c(0.025, 0.975))
```

```
  theta_quantile = quantile(theta_draw, probs = c(0.025, 0.975))

  c("theta mean" = theta_mean,"sigma mean" = sigma_mean,
    "theta 2.5% quantile" = theta_quantile[1], "theta 97.5% quantile" = theta_quantile[2],
    "sigma 2.5% quantile" = sigma_quantile[1], "sigma 97.5% quantile" = sigma_quantile[2] )
}
```

Compute school-specific posterior θ and σ

```
compute_post(school_data = school1, S = 5000)
```

```
##                 theta mean                  sigma mean
##                   9.261543                    3.906082
##   theta 2.5% quantile.2.5% theta 97.5% quantile.97.5%
##                   7.738003                   10.767281
##   sigma 2.5% quantile.2.5% sigma 97.5% quantile.97.5%
##                   3.014765                    5.161213
```

```
compute_post(school_data = school2, S = 5000)
```

```
##                 theta mean                  sigma mean
##                   6.951241                    4.406502
##   theta 2.5% quantile.2.5% theta 97.5% quantile.97.5%
##                   5.184845                    8.740655
##   sigma 2.5% quantile.2.5% sigma 97.5% quantile.97.5%
##                   3.365802                    5.842561
```

```
compute_post(school_data = school3, S = 5000)
```

```
##                 theta mean                  sigma mean
##                   7.815901                    3.750348
##   theta 2.5% quantile.2.5% theta 97.5% quantile.97.5%
##                   6.153923                    9.468380
##   sigma 2.5% quantile.2.5% sigma 97.5% quantile.97.5%
##                   2.794819                    5.156338
```

b. The posterior probability that $\theta_i < \theta_j < \theta_k$ for all six permutations {i,j,k} of {1,2,3}.

```
draw_post_theta <- function(school_data)
{
  n = length(school_data)
  ybar = mean(school_data)
  s2 = var(school_data)

  kn = k0 + n
  nun = nu0 + n
  mun = (k0 * mu0 + sum(school_data)) / kn
  sigma_n2 = (1 / nun) * (nu0 * sigma02 + (n - 1) * s2 + ((k0 * n) / kn) * (ybar - mu0)^2)

  sigma_2_draw <- 1 / rgamma(5000, nun / 2, nun * sigma_n2 / 2)
  theta_draw <- rnorm(5000, mun, sqrt(sigma_2_draw / kn))
  return(theta_draw)
}
```

```r
three_school_theta <- cbind(draw_post_theta(school1),
                            draw_post_theta(school2),
                            draw_post_theta(school3))

combination <- list(c(1,2,3),c(1,3,2),c(2,1,3),c(2,3,1),c(3,1,2),c(3,2,1))

order <- sapply(combination, function(x) { paste(x, collapse =' < ')})
probability <- sapply(combination, function(x) {
  mean(three_school_theta[, x[1] ] < three_school_theta[, x[2] ] &
         three_school_theta[, x[2] ] < three_school_theta[, x[3] ]) })

tibble(order, probability)
```

```
## # A tibble: 6 x 2
##   order     probability
##   <chr>           <dbl>
## 1 1 < 2 < 3      0.0062
## 2 1 < 3 < 2      0.0044
## 3 2 < 1 < 3      0.0756
## 4 2 < 3 < 1      0.678
## 5 3 < 1 < 2      0.0166
## 6 3 < 2 < 1      0.219
```

c. The posterior probability that $\tilde{Y}_i < \tilde{Y}_j < \tilde{Y}_k$ for all six permutations {i,j,k} of {1,2,3}.

```r
draw_post_sigma2 <- function(school_data)
{
  n = length(school_data)
  ybar = mean(school_data)
  var = var(school_data)

  kn = k0 + n
  nun = nu0 + n
  mun = (k0 * mu0 + n * ybar) / kn
  sigma_n2 = (1 / nun) * (nu0 * sigma02 + (n - 1) * var + ((k0 * n) / kn) * (ybar - mu0)^2)

  sigma_2_draw <- 1 / rgamma(5000, nun / 2, nun * sigma_n2 / 2)
  theta_draw <- rnorm(5000, mun, sqrt(sigma_2_draw / kn))
  return(sigma_2_draw)
}
```

```r
three_school_sigma2 <- cbind(draw_post_sigma2(school1),
                             draw_post_sigma2(school2),
                             draw_post_sigma2(school3))
predict_y <- list()
for (i in 1:3)
{predict_y[[i]] = rnorm(5000, three_school_theta[,i], sqrt(three_school_sigma2[,i]))}

probability <- sapply(combination, function(x) {
  mean(predict_y[[ x[1] ]] < predict_y[[ x[2] ]]&
         predict_y[[ x[2] ]] < predict_y[[ x[3] ]]) })

tibble(order, probability)
```

```
## # A tibble: 6 x 2
##   order     probability
##   <chr>           <dbl>
## 1 1 < 2 < 3       0.105
## 2 1 < 3 < 2       0.110
## 3 2 < 1 < 3       0.178
## 4 2 < 3 < 1       0.273
## 5 3 < 1 < 2       0.131
## 6 3 < 2 < 1       0.202
```

d. The posterior probability that θ_1 is bigger than both θ_2 and θ_3 is around 0.89, and the posterior probability that $\tilde{Y}_1$ is bigger than both $\tilde{Y}_2$ and $\tilde{Y}_3$ is around 0.47.

```
mean(three_school_theta[,1] > three_school_theta[,2] & three_school_theta[,1]> three_school_theta[,3])
```

```
## [1] 0.8972
```

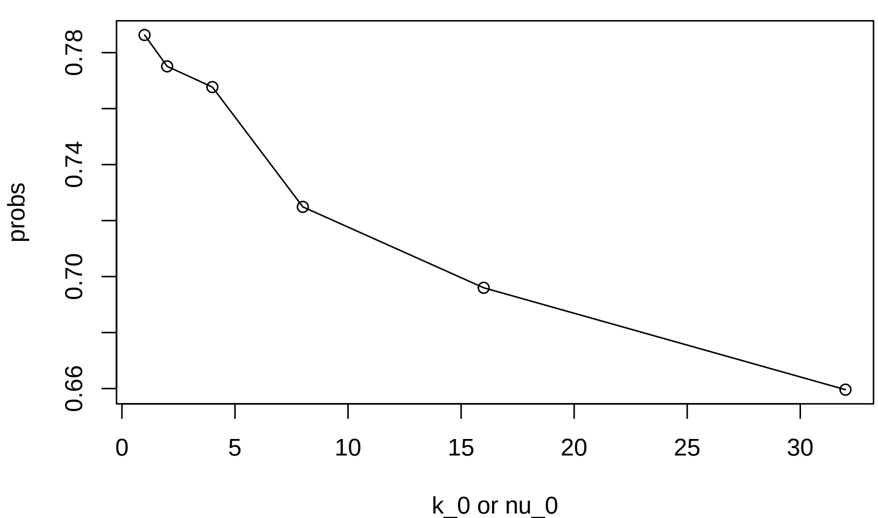
```
mean(predict_y[[1]] > predict_y[[2]] & predict_y[[1]]> predict_y[[3]])
```

```
## [1] 0.4756
```

5.2

From the plot, the posterior usually does support that $\theta_A < \theta_B$. As prior belief gets stronger, we have weaker evidence that $\theta_A < \theta_B$. If the prior opinion is very strong, the posterior might reject that $\theta_A < \theta_B$.

```
mu0 <- 75
s20 <- 100
n_A = n_B <- 16
ybar_A <- 75.2
s2_A <- 7.3^2
ybar_B <- 77.5
s2_B <- 8.1^2
k0.nu0 <- c(1, 2, 4, 8, 16, 32)
probs <- sapply(k0.nu0, function(k0.nu0)
  {
  kn_A = k0.nu0 + n_A
  nun_A = k0.nu0 + n_A
  mun_A = (k0.nu0 * mu0 + n_A * ybar_A) / kn_A
  s2n_A = (1 / nun_A) * (k0.nu0 * s20 + (n_A - 1) * s2_A + ((k0.nu0 * n_A) / kn_A) * (ybar_A - mu0)^2)
  s2_A.draw = 1 / rgamma(10000, nun_A / 2, s2n_A * nun_A / 2)
  theta_A.draw = rnorm(10000, mun_A, sqrt(s2_A.draw/kn_A))

  kn_B = k0.nu0 + n_B
  nun_B = k0.nu0 + n_B
  mun_B = (k0.nu0 * mu0 + n_B * ybar_B) / kn_B
  s2n_B = (1 / nun_B) * (k0.nu0 * s20 + (n_B - 1) * s2_B + ((k0.nu0 * n_B) / kn_B) * (ybar_B - mu0)^2)
  s2_B.draw = 1 / rgamma(10000, nun_B / 2, s2n_B * nun_B / 2)
  theta_B.draw = rnorm(10000, mun_B, sqrt(s2_B.draw/kn_B))

  mean(theta_A.draw < theta_B.draw)
})
plot(k0.nu0, probs, type = "o", xlab = "k_0 or nu_0")
```



6.1

a. θ_A and θ_B are not independent under this prior distribution.

Such a joint prior distribution is justified when we assume θ_B is proportional to θ_A, which follows Gamma distribution, and the rate also follows Gamma distribution.

b. The full conditional for θ is given as

$$
\begin{aligned}
\theta \sim Gamma(a_\theta, b_\theta), &\quad \gamma \sim Gamma(a_\gamma, b_\gamma) \\
\boldsymbol{y}_A \sim Poisson(\theta_A = \theta), &\quad \boldsymbol{y}_B \sim Poisson(\theta_B = \gamma\theta), \\
p(\theta, \mid \boldsymbol{y}_A, \boldsymbol{y}_B, \gamma) &\propto p(\theta) \times p(\gamma) \times p(\boldsymbol{y}_A \mid \theta) \times p(\boldsymbol{y}_B \mid \theta, \gamma) \\
&\propto \left(\theta^{a_\theta - 1} e^{-b_\theta \theta}\right) \times \left(\gamma^{a_\gamma - 1} e^{-b_\gamma \gamma}\right) \times \left(\prod_{i=1}^{n_A} \theta^{y_i} e^{-\theta}\right) \times \left(\prod_{j=1}^{n_B} (\gamma\theta)^{y_j} e^{-\gamma\theta}\right) \\
&= \left(\theta^{a_\theta - 1} e^{-b_\theta \theta}\right) \times \left(\gamma^{a_\gamma - 1} e^{-b_\gamma \gamma}\right) \times \left(\theta^{\sum_{i=1}^{n_A} y_i} e^{-n_A \theta}\right) \times \left((\gamma\theta)^{\sum_{j=1}^{n_B} y_j} e^{-n_B \gamma\theta}\right) \\
\text{Only care about } \theta &\propto \left(\theta^{a_\theta - 1} e^{-b_\theta \theta}\right) \times \left(\gamma^{a_\gamma - 1} e^{-b_\gamma \gamma}\right) \times \left(\theta^{n_A \bar{y}_A} e^{-n_A \theta}\right) \times \left((\gamma\theta)^{n_B \bar{y}_B} e^{-n_B \gamma\theta}\right) \\
&\propto \left(\theta^{a_\theta - 1} e^{-b_\theta \theta}\right) \times \left(\theta^{n_A \bar{y}_A} e^{-n_A \theta}\right) \times \left((\gamma\theta)^{n_B \bar{y}_B} e^{-n_B \gamma\theta}\right) \\
&\propto \theta^{a_\theta + n_A \bar{y}_A + n_B \bar{y}_B - 1} \exp\left(-(b_\theta + n_A + n_B \gamma)\theta\right) \\
&\propto \text{Gamma}\left(a_\theta + n_A \bar{y}_A + n_B \bar{y}_B, b_\theta + n_A + n_B \gamma\right) \\
\text{Or } &\propto \text{Gamma}\left(a_\theta + \sum_{i}^{n_A} y_i + \sum_{j}^{n_B} y_j, b_\theta + n_A + n_B \gamma\right)
\end{aligned}
$$

c. The full conditional for γ is given as

$$
\begin{aligned}
p(\gamma, \mid \boldsymbol{y}_A, \boldsymbol{y}_B, \theta) &\propto \left(\theta^{a_\theta - 1} e^{-b_\theta \theta}\right) \times \left(\gamma^{a_\gamma - 1} e^{-b_\gamma \gamma}\right) \times \left(\theta^{n_A \bar{y}_A} e^{-n_A \theta}\right) \times \left((\gamma\theta)^{n_B \bar{y}_B} e^{-n_B \gamma\theta}\right) \\
\text{Only care about } \gamma &\propto \left(\gamma^{a_\gamma - 1} e^{-b_\gamma \gamma}\right) \times \left((\gamma\theta)^{n_B \bar{y}_B} e^{-n_B \gamma\theta}\right) \\
&\propto \left(\gamma^{a_\gamma - 1} e^{-b_\gamma \gamma}\right) \times \left(\gamma^{n_B \bar{y}_B} e^{-n_B \gamma\theta}\right) \\
&\propto \gamma^{a_\gamma + n_B \bar{y}_B - 1} \exp\left(-(b_\gamma + n_B \theta)\gamma\right) \\
&\propto \text{Gamma}\left(a_\gamma + n_B \bar{y}_B, b_\gamma + n_B \theta\right) \\
\text{Or } &\propto \text{Gamma}\left(a_\gamma + \sum_{j}^{n_B} y_j, b_\gamma + n_B \theta\right)
\end{aligned}
$$

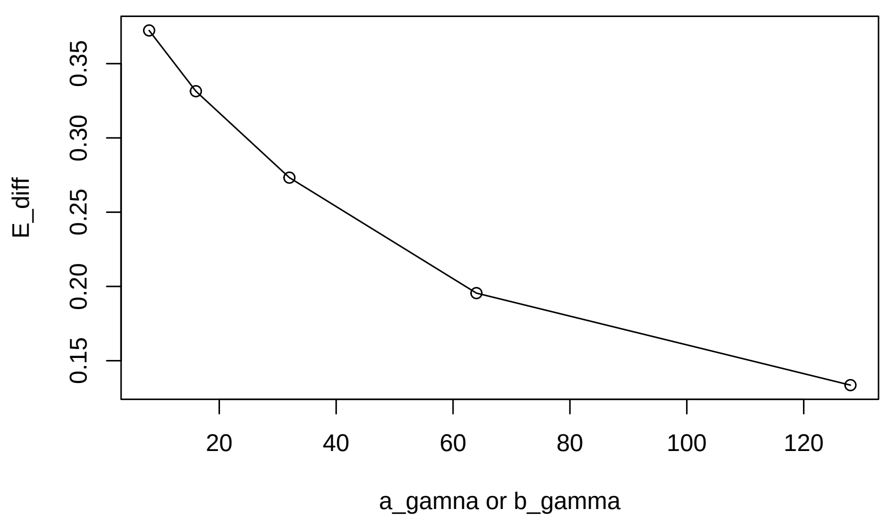
d.

```
Y_a <- scan("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/menchild30bach.dat")
Y_b <- scan("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/menchild30nobach.dat")
```

```
n_a <- length(Y_a)
n_b <- length(Y_b)
ybar_a <- mean(Y_a)
ybar_b <- mean(Y_b)
sum_a <- sum(Y_a)
sum_b <- sum(Y_b)
a_theta <- 2
b_theta <- 1
ab_gamma <- c(8, 16, 32, 64, 128)
```

```r
S <- 5000
E_diff <- sapply(ab_gamma, function(ab_gamma) {
  a_gamma = b_gamma = ab_gamma
  theta_draw = numeric(S); gamma_draw = numeric(S)
  theta = 1; gamma = 2

  for (t in 1:S) {
    gamma = rgamma(1, a_gamma + sum_b, b_gamma + n_b * theta)

    theta = rgamma(1, a_theta + sum_a + sum_b, b_theta + n_a + n_b * gamma)

    theta_draw[t] = theta
    gamma_draw[t] = gamma
  }

  theta_A = theta_draw
  theta_B = theta_draw * gamma_draw
  mean(theta_B - theta_A)
})
```

```r
plot(ab_gamma, E_diff, type = "o", xlab = "a_gamna or b_gamma")
```



As a_γ and b_γ (strength of prior belief on γ) grow larger, the difference between posterior mean of θ_A and θ_B becomes smaller.