

STA 601 Homework 3

Lingyun Shao

sep. 20, 2018

Math Problem 1

Derive the mean and variance of a Poisson distribution with parameter θ from first principles.

The pmf of $X \sim \text{Poisson}(\theta)$ is

$$\Pr(X = x) = \frac{e^{-\theta} \theta^x}{x!}$$

Mean

We know $\sum_{x=0}^{\infty} \frac{e^{-\theta} \theta^x}{x!} = 1$. By definition, we have

$$\begin{aligned} E[X] &= \sum_{x=0}^{\infty} x \Pr(X = x) \\ &= \sum_{x=1}^{\infty} x \Pr(X = x) \\ &= \sum_{x=1}^{\infty} x \frac{e^{-\theta} \theta^x}{x!} \\ &= \theta \sum_{x=1}^{\infty} \frac{e^{-\theta} \theta^{x-1}}{(x-1)!} \\ &= \theta \sum_{x=0}^{\infty} \frac{e^{-\theta} \theta^x}{x!} \\ &= \theta \end{aligned}$$

Variance

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^{\infty} x(x-1) \Pr(X = x) \\ &= \sum_{x=2}^{\infty} x(x-1) \Pr(X = x) \\ &= \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\theta} \theta^x}{x!} \\ &= \theta^2 \sum_{x=2}^{\infty} \frac{e^{-\theta} \theta^{x-2}}{(x-2)!} \\ &= \theta^2 \sum_{x=0}^{\infty} \frac{e^{-\theta} \theta^x}{x!} \\ &= \theta^2 \end{aligned}$$

So $E[X^2] = E[X(X-1) + X] = E[X(X-1)] + E[X] = \theta^2 + \theta$. Thus, we have

$$Var(X) = E[X^2] - (E[X])^2 = \theta^2 + \theta - \theta^2 = \theta$$

Math Problem 2

Derive the induced prior for θ in the Poisson model when the prior for $\phi = \log \theta$ is given by $p(\phi|n_0, t_0) = \exp(-n_0 \exp(\phi)) \exp(n_0 t_0 \phi)$

By the *change of variables formula*, we have the following induced prior for θ .

$$\begin{aligned} p(\theta|n_0, t_0) &= p(\phi|n_0, t_0) \left| \frac{d\phi}{d\theta} \right| \\ &= e^{-n_0 \exp(\phi)} e^{n_0 t_0 \phi} \left| \frac{d \log \theta}{d\theta} \right| \\ &= e^{-n_0 \exp(\log(\theta))} e^{n_0 t_0 \log(\theta)} \theta^{-1} \\ &= e^{-n_0 \theta} \theta^{n_0 t_0} \theta^{-1} \\ &= e^{-n_0 \theta} \theta^{n_0 t_0 - 1} \end{aligned}$$

3.3

Tumor counts: A cancer laboratory is estimating the rate of tumorigenesis in two strains of mice, A and B . They have tumor count data for 10 mice in strain A and 13 mice in strain B . Type A mice have been well studied, and information from other laboratories suggests that type A mice have tumor counts that are approximately Poisson-distributed with a mean of 12. Tumor count rates for type B mice are unknown, but type B mice are related to type A mice. The observed tumor counts for the two populations are

$$\begin{aligned} y_A &= (12, 9, 12, 14, 13, 13, 15, 8, 15, 6); \\ y_B &= (11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7). \end{aligned}$$

a) Find the posterior distributions, means, variances and 95% quantilebased confidence intervals for θ_A and θ_B , assuming a Poisson sampling distribution for each group and the following prior distribution:

$$\theta_A \sim \text{gamma}(120, 10), \theta_B \sim \text{gamma}(12, 1), p(\theta_A, \theta_B) = p(\theta_A) \times p(\theta_B).$$

The observed counts should be independent from each other, which means each element in y_A and y_B are respectively i.i.d. given θ_A, θ_B . Assuming that $y_A = (12, 9, 12, 14, 13, 13, 15, 8, 15, 6) = (y_1^A, \dots, y_{10}^A)$, $y_B = (11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7) = (y_1^B, \dots, y_{13}^B)$, we have the following:

$$\begin{aligned}
p(\theta_A|y_A) &\propto p(\theta_A)p(y_A|\theta_A) \\
&= p(\theta_A) \prod_{i=1}^{10} p(y_i^A|\theta_A) \\
&\propto \theta_A^{119} e^{-10\theta_A} \prod_{i=1}^{10} (e^{-\theta_A} \theta_A^{y_i^A}) \\
&= \theta_A^{119+\sum_{i=1}^{10} y_i^A} e^{-20\theta_A} \\
&= \theta_A^{236} e^{-20\theta_A}
\end{aligned}$$

We can do the similar analysis to θ_B and get $p(\theta_B|y_B) \propto \theta_B^{124} e^{-14\theta_B}$

Judging from the kernels of their posterior distribution, we can easily determine that $\theta_A|y_A \sim \text{gamma}(237, 20)$, $\theta_B|y_B \sim \text{gamma}(125, 14)$. The mean and variance of a $\text{gamma}(a, b)$ distribution are $\frac{a}{b}$ and $\frac{a}{b^2}$, so we have $E[\theta_A|y_A] = \frac{237}{20}$, $E[\theta_B|y_B] = \frac{125}{14}$ and $E[\theta_A|y_A] = \frac{237}{400}$, $E[\theta_B|y_B] = \frac{125}{196}$.

```

y.a = c(12,9,12,14,13,13,15,8,15,6) #observation of A
y.b = c(11,11,10,9,9,8,7,10,6,8,8,9,7) #observation of B
al_0.a = 120 #prior para of A
be_0.a = 10 #prior para of A
al_0.b = 12 #prior para of B
be_0.b = 1 #prior para of B
al_p.a = al_0.a + sum(y.a) #posterior para of A
be_p.a = be_0.a + length(y.a) #posterior para of A
al_p.b = al_0.b + sum(y.b) #posterior para of B
be_p.b = be_0.b + length(y.b) #posterior para of B
m_p.a = al_p.a/be_p.a #posterior mean of A
m_p.b = al_p.b/be_p.b #posterior mean of B
v_p.a = al_p.a/be_p.a^2 #posterior var of A
v_p.b = al_p.b/be_p.b^2 #posterior var of B
ci_p.a = qgamma(c(0.025, 0.975), al_p.a, be_p.a) #.95 CI of A
ci_p.b = qgamma(c(0.025, 0.975), al_p.b, be_p.b) #.95 CI of B
stat = data.frame(mean = c(m_p.a, m_p.b), var = c(v_p.a, v_p.b),
                  CI = rbind(ci_p.a, ci_p.b))
rownames(stat) = c('thetaA', 'thetaB')
colnames(stat)[3:4] = c('CI.lower', 'CI.upper')
knitr::kable(stat, digits = 2, caption = 'Posterior Means, variances and CI\'s')

```

Table 1: Posterior Means, variances and CI's

	mean	var	CI.lower	CI.upper
thetaA	11.85	0.59	10.39	13.41
thetaB	8.93	0.64	7.43	10.56

b) Compute and plot the posterior expectation of θ_B under the prior distribution $\theta_B \sim \text{gamma}(12 \times n_0, n_0)$ for each value of $n_0 \in 1, 2, \dots, 50$. Describe what sort of prior beliefs about θ_B would be necessary in order for the posterior expectation of θ_B to be close to that of θ_A .

```

n_0 = 1:50
y.b = c(11,11,10,9,9,8,7,10,6,8,8,9,7) #observation of B
al_0.b = 12*n_0 #prior para of B

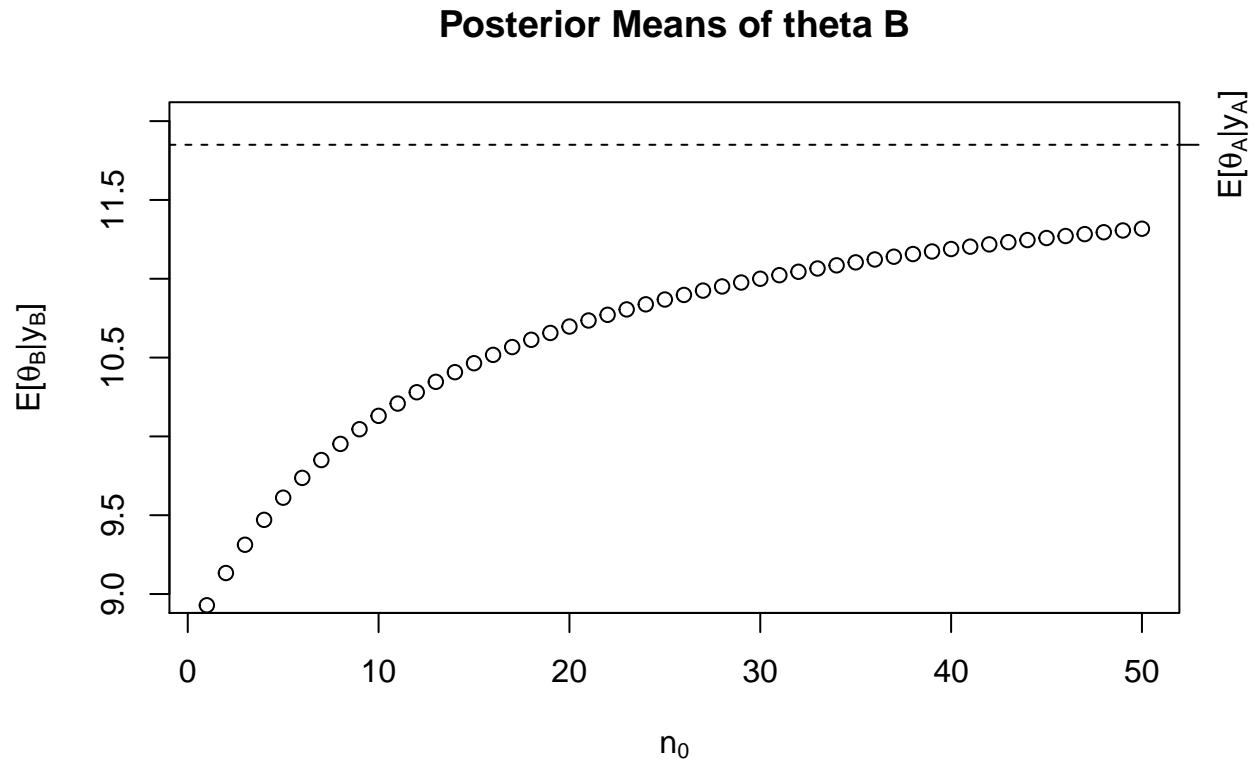
```

```

be_0.b = n_0 #prior para of B
al_p.b = al_0.b + sum(y.b) #posterior para of B
be_p.b = be_0.b + length(y.b) #posterior para of B
m_p.b = al_p.b/be_p.b #posterior mean of B

plot(n_0, m_p.b, type = 'b', ylim = c(9,12),
     xlab = expression(n[0]),
     ylab = expression(paste("E[", theta[B], "|", y[B], "]", sep = "")),
     main = 'Posterior Means of theta B')
abline(h = m_p.a, lty = 2)
axis(4, m_p.a, expression(paste("E[", theta[A], "|", y[A], "]", sep = "")))

```



We know the prior mean of θ_B is $\frac{12n_0}{n_0} = 12$ and sample mean of y_B is 8.69.

$$E[\theta_B | y_B] = \frac{12n_0 + \sum y_i^B}{n_0 + 13} = \frac{n_0}{n_0 + 13} 12 + \frac{13}{n_0 + 13} \bar{y}_B$$

In order for the posterior expectation of θ_B to be close to $E[\theta_A | y_A] = 11.85$, we need to put more weight on the prior mean, which is greater, to pull the posterior mean up a little bit. Thus, n_0 should be large, that is 50 in the given setting. Actually we can derive that when $n_0 \approx 274$, $E[\theta_B | y_B] \approx 11.85 = E[\theta_A | y_A]$.

c) Should knowledge about population A tell us anything about population B? Discuss whether or not it makes sense to have $p(\theta_A, \theta_B) = p(\theta_A) \times p(\theta_B)$.

Since we are already given the statement that type B mice are related to type A mice, knowing some information about type A mice should tell us something about type B mice. Then we should not accept that $p(\theta_A, \theta_B) = p(\theta_A) \times p(\theta_B)$ in that if the equation holds then our prior beliefs about the tumor rate among

type A mice and that of type B's are independent. When we explicitly know there is some correlation between these two populations, our prior beliefs about the tumor rates in these two types should not be independent.

3.9

Galenshore distribution: An unknown quantity Y has a Galenshore(a, θ) distribution if its density is given by

$$p(y) = \frac{2}{\Gamma(a)} \theta^{2a} y^{2a-1} e^{-\theta^2 y^2}$$

for $y > 0$, $\theta > 0$ and $a > 0$. Assume for now that a is known. For this density,

$$E[Y] = \frac{\Gamma(a + 1/2)}{\theta \Gamma(a)}, \quad E[Y^2] = \frac{a}{\theta^2}.$$

a) Identify a class of conjugate prior densities for θ . Plot a few members of this class of densities.

First of all, we notice that Galenshore distribution is actually another form of Gamma distribution in that if we let $y = t^2$ in $Gamma(a, \theta^2)$, then

$$p(y) = \frac{\theta^{2a}}{\Gamma(a)} y^a e^{-\theta^2 y}$$

$$\xrightarrow{y=t^2} p(t) = \frac{\theta^{2a}}{\Gamma(a)} t^{2a} e^{-\theta^2 t^2} \left| \frac{d(t^2)}{dt} \right| = \frac{2}{\Gamma(a)} \theta^{2a} t^{2a-1} e^{-\theta^2 t^2}$$

which is exactly the same form as Galenshore distribution.

Say our prior is $p(\theta)$, then the posterior

$$p(\theta|y) \propto p(\theta)p(y|\theta) = p(\theta)\theta^{2a} \exp\{-\theta^2 y^2\}$$

To get a class of conjugate prior densities, we have to make sure the kernel of the prior pdf is identical in form to the posterior kernel, which means there has to be $\theta^{c_1} \exp\{-c_2 \theta^2\}$ in the prior pdf. We notice that a galenshore distribution with parameters of (a_0, θ_0) itself has a kernel of this form. Then $c_1 = 2a_0 - 1$, $c_2 = \theta_0^2$, and the prior should be

$$p(\theta) = \frac{2}{\Gamma(a_0)} \theta_0^{2a_0} \theta^{2a_0-1} e^{-\theta_0^2 \theta^2}$$

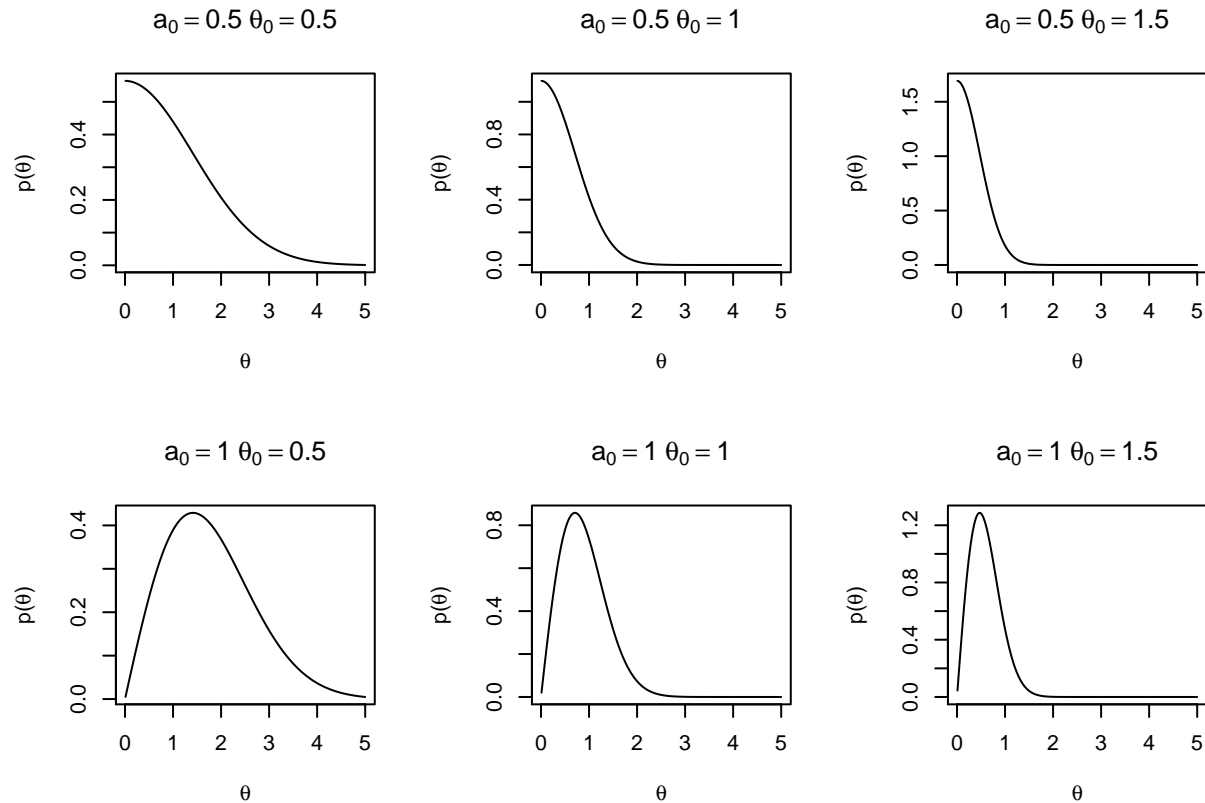
```
dgal = function(x, a, theta) {
  (2/gamma(a)) * (theta^(2*a)) * x^{2*a-1} *exp(-theta^2*x^2)
}

theta = seq(0.01, 5, by = 0.01)
a0 = c(0.5,1)
theta0 = seq(0.5, 1.5, by = 0.5)
par(mfrow=c(length(a0), length(theta0)))
for (i in 1:length(a0)) {
  for (j in 1:length(theta0)) {
    plot(theta, dgal(theta, a0[i], theta0[j]), type = 'l',
      xlab = expression(theta), ylab = expression(
        paste('p(', theta, ')', sep = '')),
```

```

main = bquote(a[0] == .(a0[i]) ~ theta[0] == .(theta0[j]))
}
}

```



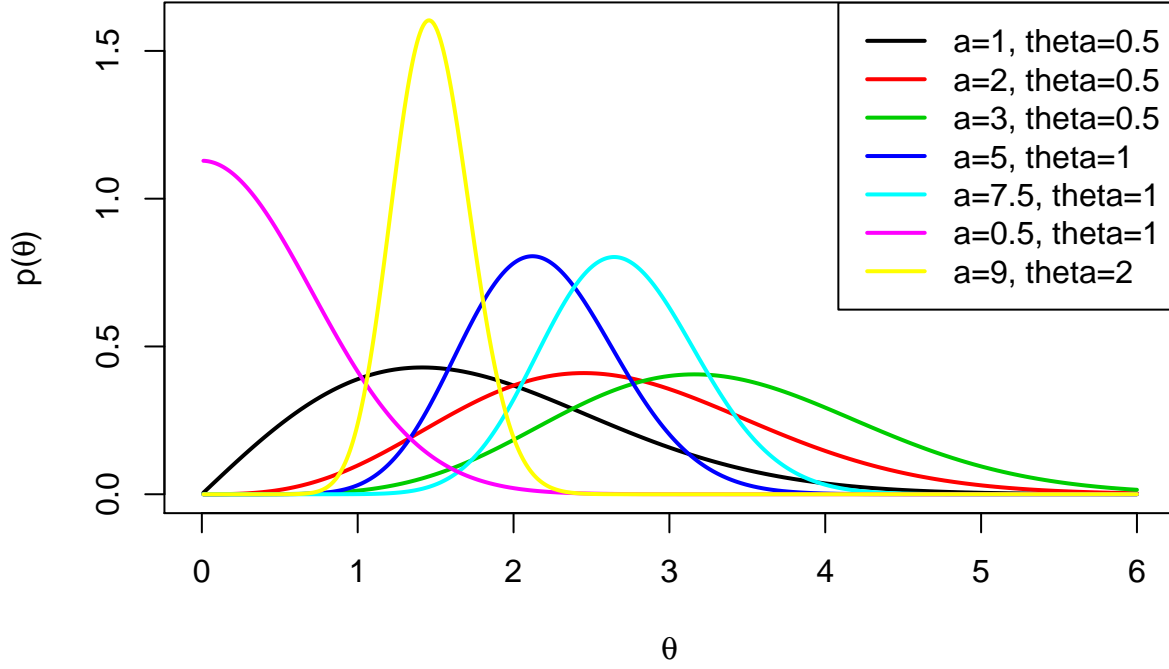
```

theta = seq(0.01, 6, by = 0.01)
a0 = c(1, 2, 3, 5, 7.5, 0.5, 9)
theta0 = c(rep(0.5, 3), rep(1, 3), 2)
par(mfrow=c(1,1))
plot(NA, xlim=c(0,6), ylim = c(0,1.6), xlab = expression(theta), ylab = expression(
  paste('p(', theta, ')', sep = '')),
  main = 'pdf\'s of Galenshore distributions')
legend = NULL
for(i in 1:length(a0)) {
  lines(theta, dgal(theta, a0[i], theta0[i])
    , col = i, lwd = 2)
  legend = c(legend, (paste('a=', a0[i], ', ', 'theta', '=', theta0[i], sep = '')))
}

legend('topright', legend = legend, lwd = 2, col=1:7)

```

pdf's of Galenshore distributions



b) Let $Y_1, \dots, Y_n \sim \text{i.i.d. Galenshore}(a, \theta)$. Find the posterior distribution of θ given Y_1, \dots, Y_n , using a prior from your conjugate class.

Say the prior is $\text{Galenshore}(a_0, \theta_0)$, then the posterior distribution given Y_1, \dots, Y_n is

$$\begin{aligned}
 p(\theta|y_1, \dots, y_n) &\propto p(\theta) \prod_{i=1}^n p(y_i|\theta) \\
 &\propto \theta^{2a_0-1} \exp\{-\theta_0^2 \theta^2\} \theta^{2na} \exp\left\{-\theta^2 \sum_{i=1}^n y_i^2\right\} \\
 &= \theta^{2(a_0+na)-1} \exp\left\{-\theta^2(\theta_0^2 + \sum_{i=1}^n y_i^2)\right\}
 \end{aligned}$$

From the kernel, we know that the posterior $p(\theta|y_1, \dots, y_n)$ should be $\text{Galenshore}(a_0 + na, \sqrt{\theta_0^2 + \sum_{i=1}^n y_i^2})$.

c) Write down $p(\theta_a|Y_1, \dots, Y_n)/p(\theta_b|Y_1, \dots, Y_n)$ and simplify. Identify a sufficient statistic.

$$\begin{aligned}
 \frac{p(\theta_a|y_1, \dots, y_n)}{p(\theta_b|y_1, \dots, y_n)} &= \frac{\theta_a^{2(a_0+na)-1} \exp\{-\theta_a^2(\theta_0^2 + \sum_{i=1}^n y_i^2)\}}{\theta_b^{2(a_0+na)-1} \exp\{-\theta_b^2(\theta_0^2 + \sum_{i=1}^n y_i^2)\}} \\
 &= \left(\frac{\theta_a}{\theta_b}\right)^{2(a_0+na)-1} \exp\left\{-(\theta_a^2 - \theta_b^2)(\theta_0^2 + \sum_{i=1}^n y_i^2)\right\}
 \end{aligned} \tag{1}$$

$p(\theta_a|y_1, \dots, y_n)/p(\theta_b|y_1, \dots, y_n)$ can be simplified as (1). As is shown in (1), the ratio of probabilities at θ_a and θ_b depends on y_i 's only through $\sum_{i=1}^n y_i^2$, which means that $\sum_{i=1}^n y_i^2$ contains all the information from y_i 's regarding θ . Thus, $\sum_{i=1}^n y_i^2$ is a sufficient statistic.

d) Determine $E[\theta|y_1, \dots, y_n]$.

Since we already know the form of a Galenshore distribution's expectation and that the posterior follows a $\text{Galenshore}(a_0 + na, \sqrt{\theta_0^2 + \sum_{i=1}^n y_i^2})$, so the posterior expectation

$$E[\theta|y_1, \dots, y_n] = \frac{\Gamma(a_0 + na + \frac{1}{2})}{\sqrt{\theta_0^2 + \sum_{i=1}^n y_i^2} \Gamma(a_0 + na)}$$

e) Determine the form of the posterior predictive density $p(\tilde{y}|y_1, \dots, y_n)$.

$$\begin{aligned} p(\tilde{y}|y_1, \dots, y_n) &= \int_0^\infty p(\tilde{y}|\theta, y_1, \dots, y_n) p(\theta|y_1, \dots, y_n) d\theta \\ &= \int_0^\infty p(\tilde{y}|\theta) p(\theta|y_1, \dots, y_n) d\theta \quad (\text{By conditional independence}) \\ &= \int_0^\infty \frac{2}{\Gamma(a)} \theta^{2a} \tilde{y}^{2a-1} \exp\{-\theta^2 \tilde{y}^2\} \frac{2}{\Gamma(a_0 + na)} (\theta_0^2 + \sum_{i=1}^n y_i^2)^{a_0+an} \theta^{2(a_0+na)-1} \exp\left\{-\theta^2(\theta_0^2 + \sum_{i=1}^n y_i^2)\right\} d\theta \\ &= \frac{2}{\Gamma(a)} \tilde{y}^{2a-1} \frac{2}{\Gamma(a_0 + na)} (\theta_0^2 + \sum_{i=1}^n y_i^2)^{a_0+an} \int_0^\infty \theta^{2(a_0+(n+1)a)-1} \exp\left\{-\theta^2(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2)\right\} d\theta \\ &= \frac{4}{\Gamma(a)\Gamma(a_0 + na)} \tilde{y}^{2a-1} (\theta_0^2 + \sum_{i=1}^n y_i^2)^{a_0+an} \frac{\Gamma(a_0 + (n+1)a)}{2(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2)^{(a_0+(n+1)a)}} \\ &\quad \times \underbrace{\int_0^\infty \frac{2(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2)^{(a_0+(n+1)a)}}{\Gamma(a_0 + (n+1)a)} \theta^{2(a_0+(n+1)a)-1} \exp\left\{-\theta^2(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2)\right\} d\theta}_{\text{integral of pdf is 1}} \\ &= \frac{4}{\Gamma(a)\Gamma(a_0 + na)} \tilde{y}^{2a-1} (\theta_0^2 + \sum_{i=1}^n y_i^2)^{a_0+an} \frac{\Gamma(a_0 + (n+1)a)}{2(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2)^{(a_0+(n+1)a)}} \\ &= \frac{2\Gamma(a_0 + (n+1)a)}{\Gamma(a)\Gamma(a_0 + na)} \frac{(\theta_0^2 + \sum_{i=1}^n y_i^2)^{a_0+an}}{(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2)^{(a_0+(n+1)a)}} \tilde{y}^{2a-1} \end{aligned}$$

4.1

Posterior comparisons: Reconsider the sample survey in Exercise 3.1. Suppose you are interested in comparing the rate of support in that county to the rate in another county. Suppose that a survey of sample size 50 was done in the second county, and the total number of people in the sample who supported the policy was 30. Identify the posterior distribution of θ_2 assuming a uniform prior. Sample 5,000 values of each of θ_1 and θ_2 from their posterior distributions and estimate $Pr(\theta_1 < \theta_2 | \text{the data and prior})$.

$p(\theta_2|Y_2) \propto p(\theta_2)p(Y_2|\theta_2) = 1 \times \theta_2^{30}(1 - \theta_2)^{20}$, so the posterior distribution of θ_2 is $\text{beta}(31, 21)$. Recall that the posterior of θ_1 is $\text{beta}(58, 44)$.


```
n = 5000
set.seed(1)
th1 = rbeta(n, 58, 44)
th2 = rbeta(n, 31, 21)
(est = mean(th1 < th2))
```

```
## [1] 0.637
```

The estimated value of $Pr(\theta_1 < \theta_2 | \text{the data and prior})$ is 0.637 given the random seed being 1.

4.2

Tumor count comparisons: Reconsider the tumor count data in Exercise 3.3:

- a) For the prior distribution given in part a) of that exercise, obtain $Pr(\theta_B < \theta_A | y_A, y_B)$ via Monte Carlo sampling.

According to 3.3, $\theta_A | y_A \sim \text{gamma}(237, 20)$ and $\theta_B | y_B \sim \text{gamma}(125, 14)$

```
y.a = c(12,9,12,14,13,13,15,8,15,6) #observation of A
y.b = c(11,11,10,9,9,8,7,10,6,8,8,9,7) #observation of B
al_0.a = 120 #prior para of A
be_0.a = 10 #prior para of A
al_0.b = 12 #prior para of B
be_0.b = 1 #prior para of B
al_p.a = al_0.a + sum(y.a) #posterior para of A
be_p.a = be_0.a + length(y.a) #posterior para of A
al_p.b = al_0.b + sum(y.b) #posterior para of B
be_p.b = be_0.b + length(y.b) #posterior para of B
n = 10000
set.seed(1)
th_a = rgamma(n, al_p.a, be_p.a)
th_b = rgamma(n, al_p.b, be_p.b)
(pr = mean(th_b < th_a))
```

```
## [1] 0.9944
```

The estimated value of $Pr(\theta_B < \theta_A | y_A, y_B)$ is 0.9944 given the random seed being 1.

- b) For a range of values of n_0 , obtain $Pr(\theta_B < \theta_A | y_A, y_B)$ for $\theta_A \sim \text{gamma}(120, 10)$ and $\theta_B \sim \text{gamma}(12 \times n_0, n_0)$. Describe how sensitive the conclusions about the event $\{\theta_B < \theta_A\}$ are to the prior distribution on θ_B .

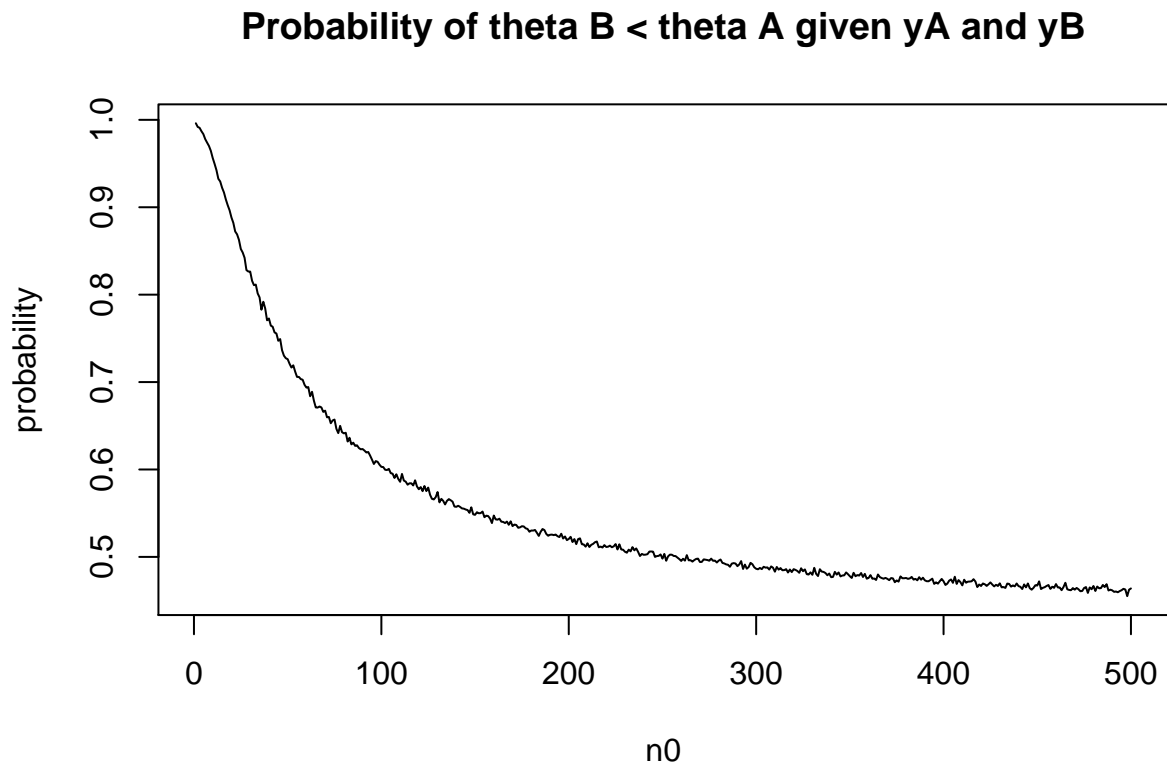
```
n = 10000
n_0 = 1:10
set.seed(1)
al_0.b = 12*n_0 #prior para of B
be_0.b = n_0 #prior para of B
al_p.b = al_0.b + sum(y.b) #posterior para of B
be_p.b = be_0.b + length(y.b) #posterior para of B
th_a = rgamma(n, al_p.a, be_p.a)
th_b = sapply(n_0, function(x) rgamma(n, al_p.b[x], be_p.b[x]))
(pr = colMeans(th_b < th_a)) #prob of thetab < thetaa
```

```
## [1] 0.9944 0.9934 0.9902 0.9894 0.9843 0.9803 0.9742 0.9644 0.9640 0.9541
```

```

n_0 = 1:500
al_0.b = 12*n_0 #prior para of B
be_0.b = n_0 #prior para of B
al_p.b = al_0.b + sum(y.b) #posterior para of B
be_p.b = be_0.b + length(y.b) #posterior para of B
th_a = rgamma(n, al_p.a, be_p.a)
th_b = sapply(n_0, function(x) rgamma(n, al_p.b[x], be_p.b[x]))
pr = colMeans(th_b < th_a)
plot(n_0, pr, type = 'l', xlab = 'n0', ylab = 'probability',
     main = 'Probability of theta B < theta A given yA and yB')

```



I first chose n_0 being from 1 to 10, we can see from the result that all probabilities are greater than 0.95, we may conclude that the event is not sensitive to prior distribution on θ_b .

However, when I chose n_0 being from 1 to 500, there showed a bigger picture of this. As is shown in the plot, the probability decreases as n_0 grows but the decreasing rate is diminishing. So we know that the probability of event $\{\theta_B < \theta_A\}$ shrinks very fast as the beginning, then it get slower and slower. When it comes to the sensitivity of the conclusion of event $\{\theta_B < \theta_A\}$, we observe that the probability gets to 0.5 approximately at $n_0 = 250$. Since we are more likely to believe $\{\theta_B < \theta_A\}$ until n_0 grows up to 250 in favor of the opposite, we may reach that the conclusions about the event is not very sensitive to the prior distribution on θ_B .

- c) Repeat parts a) and b), replacing the event $\{\theta_B < \theta_A\}$ with the event $\{\tilde{Y}_B < \tilde{Y}_A\}$, where \tilde{Y}_A and \tilde{Y}_B are samples from the posterior predictive distribution.

```

# repeat a
n = 10000

```

```

set.seed(1)
al_0.b = 12 #prior para of B
be_0.b = 1 #prior para of B
al_p.b = al_0.b + sum(y.b) #posterior para of B
be_p.b = be_0.b + length(y.b) #posterior para of B
th_a = rgamma(n, al_p.a, be_p.a)
th_b = rgamma(n, al_p.b, be_p.b)
y_a = sapply(th_a, function(x) rpois(1, x))
y_b = sapply(th_b, function(x) rpois(1, x))
(pr = mean(y_b < y_a)) #prob of thetab < thetaa

```

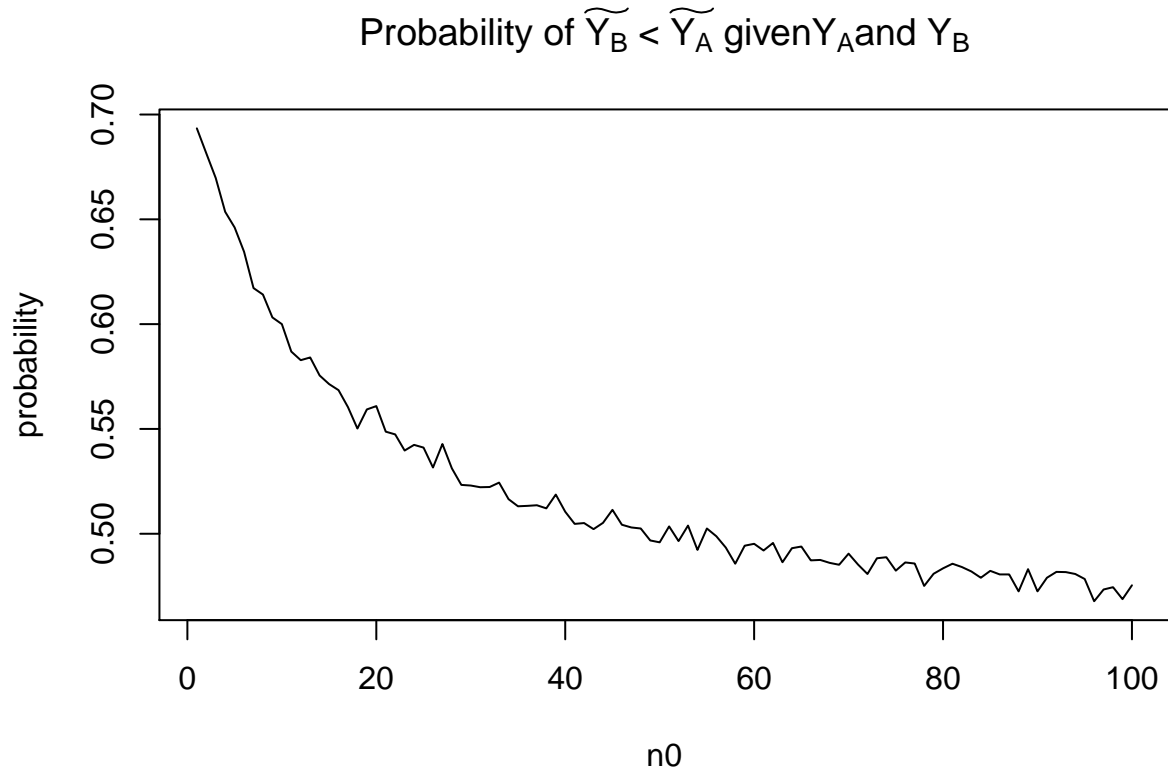
```
## [1] 0.7003
```

The estimated value of $Pr(\tilde{Y}_B < \tilde{Y}_A | y_A, y_B)$ is 0.7003 given the random seed being 1

```

#repeat b
set.seed(1)
n_0 = 1:100
al_0.b = 12*n_0 #prior para of B
be_0.b = n_0 #prior para of B
al_p.b = al_0.b + sum(y.b) #posterior para of B
be_p.b = be_0.b + length(y.b) #posterior para of B
th_a = rgamma(n, al_p.a, be_p.a)
th_b = sapply(n_0, function(x) rgamma(n, al_p.b[x], be_p.b[x]))
y_a = sapply(th_a, function(x) rpois(1, x))
y_b = apply(th_b, MARGIN = 1:2, function(x) rpois(1, x))
pr = colMeans(y_b < y_a)
plot(n_0, pr, type = 'l', xlab = 'n0', ylab = 'probability',
     main = expression(paste('Probability of ',
                             widetilde(Y[B]), ' < ', widetilde(Y[A]),
                             ' given', Y[A], 'and ', Y[B], seq=''))) )

```



I chose n_0 being from 1 to 100. We can see from the result that probabilities shrinks to 0.5 at about $n_0 = 50$. Therefore we know that the event $\{\widetilde{Y}_B < \widetilde{Y}_A\}$ is more sensitive to prior distribution on θ_b than the event in b), $\{\theta_B < \theta_A\}$.