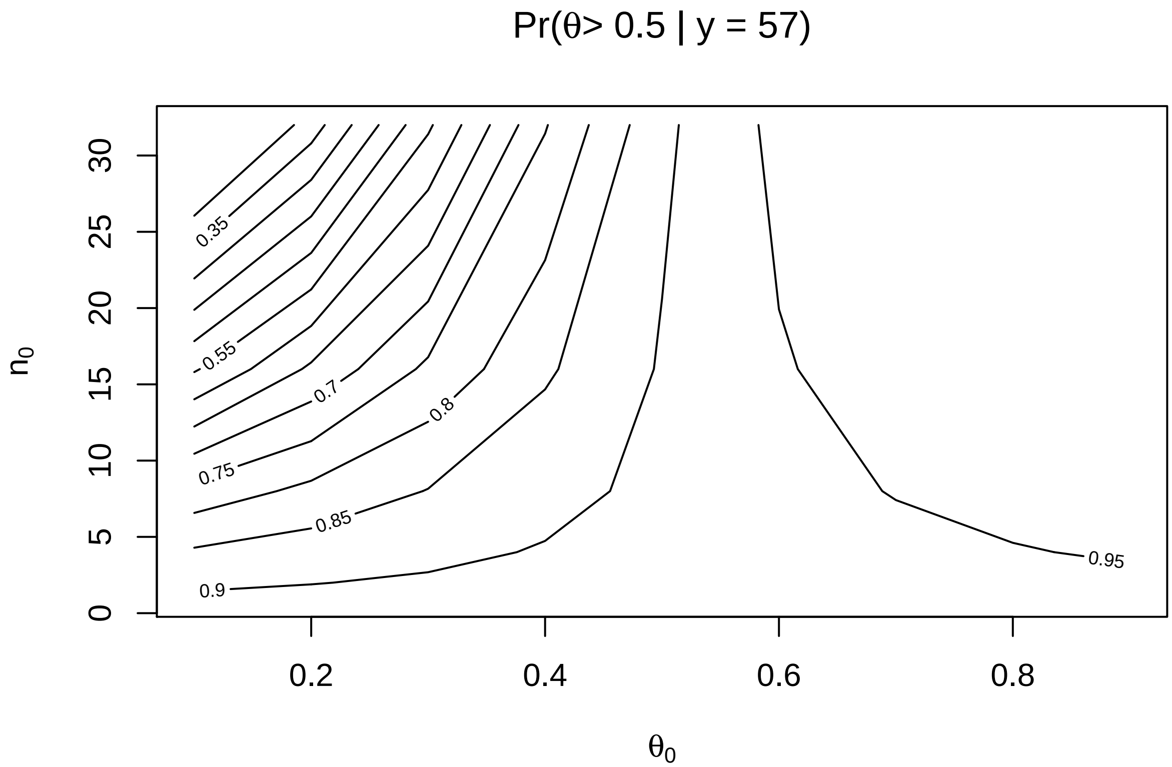
Homework 2 Solutions

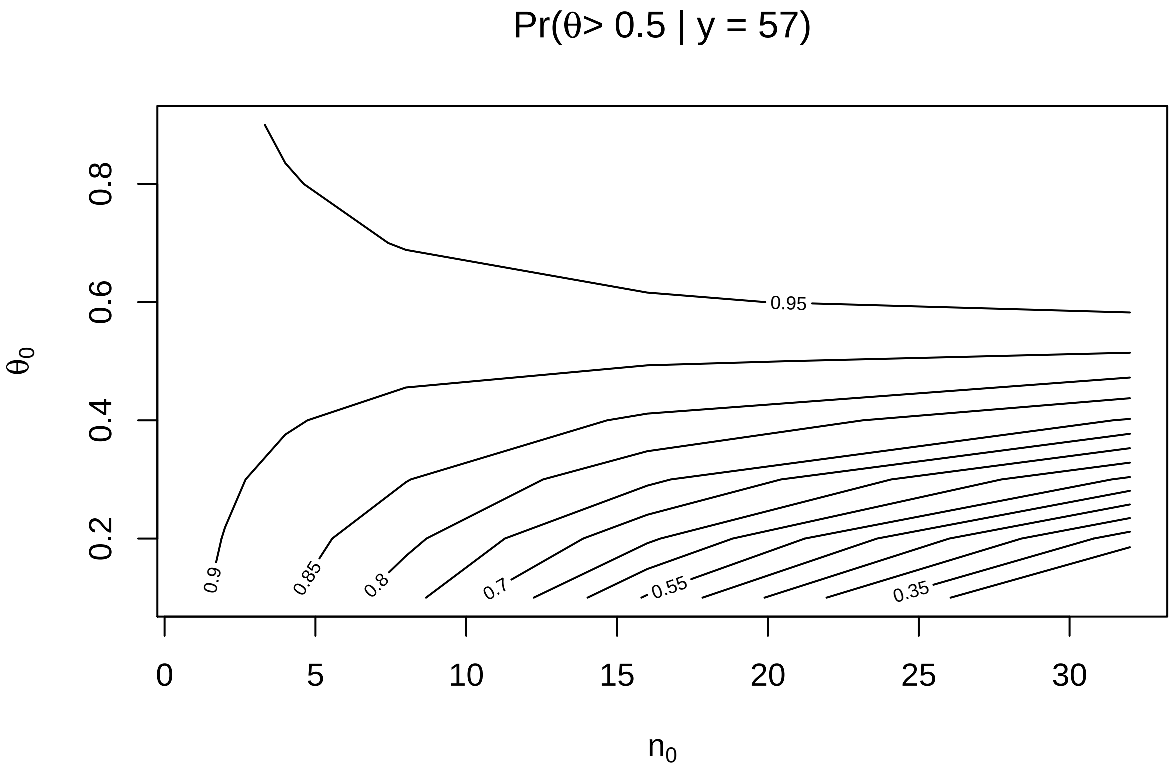
Problem 1

```r
y <- 57
n <- 100

theta.0 <- seq(0.1, 0.9, by = 0.1)
n.0 <- 2^seq(0, 5, by = 1)

post.means <- matrix(NA,
                     nrow = length(theta.0),
                     ncol = length(n.0))

for (i in 1:length(theta.0)) {
  for (j in 1:length(n.0)) {
    a.0.ij <- theta.0[i]*n.0[j]
    b.0.ij <- (1 - theta.0[i])*n.0[j]
    a.post.ij <- a.0.ij + y
    b.post.ij <- b.0.ij + n - y

    post.means[i, j] <- 1 - pbeta(0.5, a.post.ij, b.post.ij)
  }
}
```

```r
contour(theta.0, n.0, post.means,
        xlab = expression(theta[0]),
        ylab = expression(n[0]),
        levels = seq(0.3, 1, by = 0.05),
        main = expression(paste("Pr(", theta, "> 0.5 | y = 57)", sep = "")))
```




```
contour(n.0, theta.0, t(post.means),
        ylab = expression(theta[0]),
        xlab = expression(n[0]),
        levels = seq(0.3, 1, by = 0.05),
        main = expression(paste("Pr(", theta, "> 0.5 | y = 57)", sep = "")))
```



The figure shows that the posterior probability that θ exceeds 0.5 given the data depends on what our prior beliefs are, specifically what our prior belief for θ_0 is and the strength of that belief or "prior sample size," represented by n_0.

We can see that the posterior mean is not very sensitive to our prior beliefs about θ. Unless we *strongly* believe that θ is close to 0, i.e. n_0 is large and θ_0 is small, the posterior probability $\Pr(\theta > 0.5|y = 57)$ exceeds 0.5, i.e. we will conclude θ likely exceeds 0.5 given the data.

Problem 2

Part a.

The posterior distributions are $\theta_A|\boldsymbol{y}_A \sim \text{gamma}\left(120 + \sum_{i=1}^{10} y_{Ai} = 237, 10 + 10 = 20\right)$ and $\theta_B|\boldsymbol{y}_B \sim \text{gamma}\left(12 + \sum_{i=1}^{13} y_{Bi} = 125, 1 + 13 = 14\right)$.

```
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.0.a <- 120
b.0.a <- 10

a.0.b <- 12
b.0.b <- 1

a.post.a <- a.0.a + sum(y.a)
b.post.a <- b.0.a + length(y.a)
```

```r
a.post.b <- a.0.b + sum(y.b)
b.post.b <- b.0.b + length(y.b)
```

```r
e.post.a <- a.post.a/b.post.a
e.post.b <- a.post.b/b.post.b

v.post.a <- a.post.a/b.post.a^2
v.post.b <- a.post.b/b.post.b^2

ci.post.a <- qgamma(c(0.025, 0.975), a.post.a, b.post.a)
ci.post.b <- qgamma(c(0.025, 0.975), a.post.b, b.post.b)
```

The posterior means are $\mathbb{E}[\theta_A|\boldsymbol{y}_A] = 11.85$ and $\mathbb{E}[\theta_B|\boldsymbol{y}_B] = 8.93$.

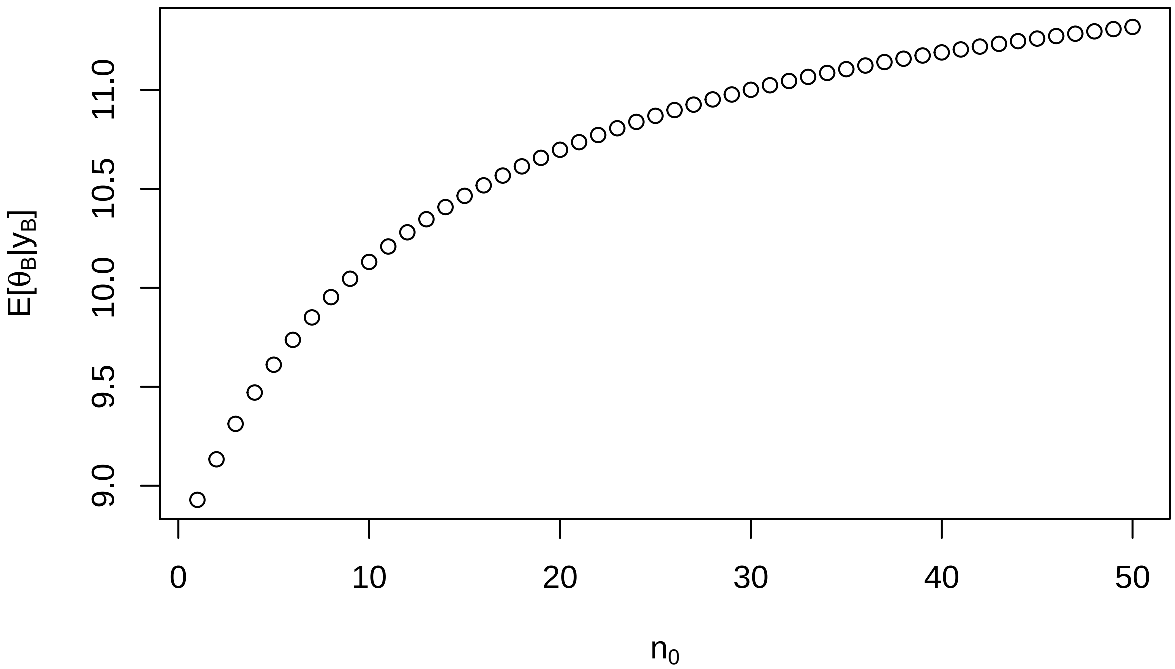
The posterior variances are $\mathbb{V}[\theta_A|\boldsymbol{y}_A] = 0.59$ and $\mathbb{V}[\theta_B|\boldsymbol{y}_B] = 0.64$.

The posterior 95% confidence intervals are $(10.39, 13.41)$ and $(7.43, 10.56)$ for θ_A and θ_B, respectively.

Part b.

```r
n.0 <- seq(1, 50)
a.0.b <- 12*n.0
b.0.b <- n.0

a.post.b <- a.0.b + sum(y.b)
b.post.b <- b.0.b + length(y.b)

e.post.b <- a.post.b/b.post.b
```

```r
plot(n.0,
     e.post.b,
     xlab = expression(n[0]),
     ylab = expression(paste("E[", theta[B], "|", y[B], "]", sep = "")))
abline(h = e.post.a, lty = 2)
```



Given that we believe that the prior mean of θ_B is 12, we would need to have very strong prior beliefs about θ_B, corresponding to $n_0 \approx 50$, in order for the posterior expectation of θ_B to be close to that of θ_A. In fact, we would need $n_0 \approx 274$ for the posterior expectation of θ_B to be equal to that of θ_A.

Part c.

Based on the information that type B mice are related to type A mice, knowledge about population A should tell us something about population B. If knowledge about type A mice should tell us something about type B mice, then it does not make sense to have $p\left(\theta_A, \theta_B\right) = p\left(\theta_A\right) p\left(\theta_B\right)$ because $p\left(\theta_A, \theta_B\right) = p\left(\theta_A\right) p\left(\theta_B\right)$ indicates that our prior beliefs about the tumor rate among type A mice are independent of our prior beliefs about the tumor rate among type B mice. If we know the two types of mice are related, our prior beliefs about the tumor rates in type A and B mice should *not* be independent.

Problem 3

$$\begin{aligned}
p\left(y\right) &= \frac{2}{\Gamma\left(a\right)}\theta^{2a}y^{2a-1}\exp\left\{-\theta^2 y^2\right\} \\
\mathbb{E}\left[Y\right] &= \frac{\Gamma\left(a+\frac{1}{2}\right)}{\theta\Gamma\left(a\right)} \\
\mathbb{E}\left[Y^2\right] &= \frac{a}{\theta^2}
\end{aligned}$$

for $y, \theta, a > 0$.

Part a.

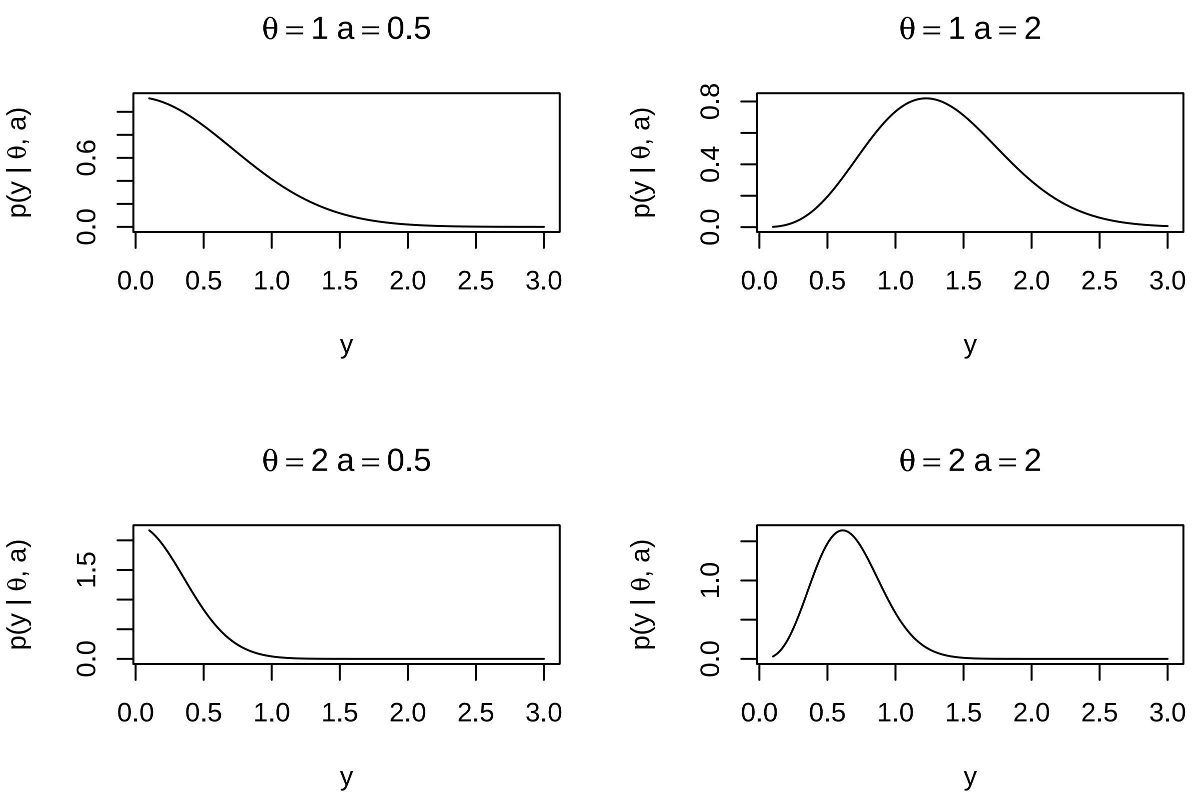
Our posterior will have the form

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

Therefore, our conjugate class of densities will have to include terms like $\theta^{c_1}\exp\left\{-c_2\theta^2\right\}$. A galenshore distribution with parameters a_0 and θ_0 has this form, with $c_1 = 2a_0 - 1$ and $c_2 = \theta_0^2$.

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

```r
y <- seq(0.1, 3, by = 0.01)

thetas <- c(1, 2)
as <- c(0.5, 2)

par(mfrow = c(2, 2))
for (i in 1:length(thetas)) {
  for (j in 1:length(as)) {
    plot(y, dgal(y, thetas[i], as[j]), type = "l",
         xlab = "y", ylab = expression(paste("p(y | ", theta, ", a)", sep = "")),
         main = bquote(theta == .(thetas[i]) ~ a == .(as[j])))
  }
}
```



Part b.

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

The posterior distribution of θ will be galenshore($na + a_0$, $\sqrt{\theta_0^2 + \sum_{i=1}^n y_i^2}$).

Part c.

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

We can see that the probability density at θ_a relative to θ_b depends on $y_1, \ldots, y_n$ only through $\sum_{i=1}^n y_i^2$. We can interpret this as meaning that $\sum_{i=1}^n y_i^2$ contains all of the information about θ available from the data, and we can say that $\sum_{i=1}^n y_i^2$ is a *sufficient statistic* for θ and $p\left(y_1, \ldots, y_n | \theta\right)$.

Part d.

Given the posterior distribution for θ given the data found in Part b. and the provided formula for the expectation of a galenshore-distributed random variable, we have

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

Part e.

$$\begin{aligned}
p\left(\tilde{y} | y_1, \ldots, y_n\right) &= \underbrace{\int_0^\infty p\left(\tilde{y} | \theta, y_1, \ldots, y_n\right) p\left(\theta | y_1, \ldots, y_n\right) d\theta}_{\text{By Bayes' rule, definition of marginal dist.}} \\
&= \underbrace{\int_0^\infty p\left(\tilde{y} | \theta\right) p\left(\theta | y_1, \ldots, y_n\right) d\theta}_{\text{By conditional indep. of } \tilde{y}, y_1, \ldots, y_n \text{ given } \theta} \\
&= \underbrace{\int_0^\infty \frac{2}{\Gamma\left(a\right)} \theta^{2a} \tilde{y}^{2a-1} \exp\left\{-\theta^2 \tilde{y}^2\right\} \frac{2}{\Gamma\left(na + a_0\right)} \left(\theta_0^2 + \sum_{i=1}^n y_i^2\right)^{(an+a_0)} \theta^{2(an+a_0)-1} \exp\left\{-\left(\theta_0^2 + \sum_{i=1}^n y_i^2\right)\theta^2\right\} d}_{\text{Plugging in likelihood of } \tilde{y} \text{ given } \theta \text{ and posterior of } \theta \text{ given } y_1, \ldots, y_n} \\
&= \underbrace{\frac{2}{\Gamma\left(a\right)} \tilde{y}^{2a-1} \frac{2\left(\theta_0^2 + \sum_{i=1}^n y_i^2\right)^{(an+a_0)}}{\Gamma\left(na + a_0\right)} \int_0^\infty \theta^{2(a(n+1)+a_0)-1} \exp\left\{-\left(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2\right)\theta^2\right\} d\theta}_{\text{Pulling constants (don't dep. on } \theta \text{) out of integral}} \\
&= \underbrace{\frac{2}{\Gamma\left(a\right)} \tilde{y}^{2a-1} \frac{2\left(\theta_0^2 + \sum_{i=1}^n y_i^2\right)^{(an+a_0)}}{\Gamma\left(na + a_0\right)} \frac{\Gamma\left(a\left(n+1\right) + a_0\right)}{2} \left(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2\right)^{2(a(n+1)+a_0)}}_{\text{Apply kernel trick - recognize } \int_0^\infty \frac{2}{\Gamma(d)} c^{2d} \theta^{2d-1} \exp\left\{-c^2\theta^2\right\} d\theta = 1} \\
&= \underbrace{\frac{2\Gamma\left(a\left(n+1\right) + a_0\right)}{\Gamma\left(a\right)\Gamma\left(na + a_0\right)} \tilde{y}^{2a-1} \left(\theta_0^2 + \sum_{i=1}^n y_i^2\right)^{(an+a_0)} \left(\theta_0^2 + \sum_{i=1}^n y_i^2 + \tilde{y}^2\right)^{2(a(n+1)+a_0)}}_{\text{Cancel some 2's and rearrange}}.
\end{aligned}$$