Lab 04

Jose Pliego San Martin

1 Mixture of conjugate priors

Let $Y|\lambda \sim \text{Poisson}(\lambda)$. Consider a prior distribution for λ given by a mixture of a $\text{gamma}(\alpha_1, \beta_1)$ and a $\text{gamma}(\alpha_2, \beta_2)$ distribution.

a

Identify the mixture components of the posterior distribution $p(\lambda|y)$.

$$\begin{aligned}
p(\lambda) &= wp_1(\lambda) + (1-w)p_2(\lambda), \\
p_1(\lambda) &= \text{dgamma}(\lambda, \alpha_1, \beta_1), \\
p_2(\lambda) &= \text{dgamma}(\lambda, \alpha_2, \beta_2), \\
p(y|\lambda) &= \lambda^y e^{-\lambda}/y!.
\end{aligned}$$

The posterior distribution is proportional to

$$\begin{aligned}
p(\lambda|y) &\propto p(\lambda)p(y|\lambda) \\
&\propto \left(wp_1(\lambda) + (1-w)p_2(\lambda)\right)p(y|\lambda) \\
&\propto wp_1(\lambda)p(y|\lambda) + (1-w)p_2(\lambda)p(y|\lambda).
\end{aligned}$$

By conjugacy,

$$\begin{aligned}
p_1(\lambda)p(y|\lambda) &\propto \text{dgamma}(\lambda, y+\alpha_1, 1+\beta_1), \\
p_2(\lambda)p(y|\lambda) &\propto \text{dgamma}(\lambda, y+\alpha_2, 1+\beta_2),
\end{aligned}$$

so the posterior distribution is a mixture of a $\text{gamma}(y+\alpha_1, 1+\beta_1)$ and a $\text{gamma}(y+\alpha_2, 1+\beta_2)$ distribution.

b

What are the weights of the posterior distribution? How can you interpret them?

To get the posterior weights, we need to look at the normalizing constant.

$$
\begin{aligned}
\int p(y|\lambda)p(\lambda)\mathrm{d}\lambda &= w\int p(y|\lambda)p_1(\lambda)\mathrm{d}\lambda + (1-w)\int p(y|\lambda)p_2(\lambda)\mathrm{d}\lambda \\
&= wp_1(y) + (1-w)p_2(y),
\end{aligned}
$$

where $p_1(y)$ and $p_2(y)$ are the prior predictive distributions induced by each of the mixture components. As seen in lecture, $p_j(y) = \text{dnegbinom}(y, \alpha_j, \beta_j)$.

Putting everything together, we get

$$
\begin{aligned}
p(\lambda|y) &= \frac{wp_1(\lambda)p(y|\lambda) + (1-w)p_2(\lambda)p(y|\lambda)}{wp_1(y) + (1-w)p_2(y)} \\
&= \frac{wp_1(y)\frac{p_1(\lambda)p(y|\lambda)}{p_1(y)} + (1-w)p_2(y)\frac{p_2(\lambda)p(y|\lambda)}{p_2(y)}}{wp_1(y) + (1-w)p_2(y)} \\
&= \frac{wp_1(y)}{wp_1(y) + (1-w)p_2(y)}p_1(\lambda|y) + \frac{(1-w)p_2(y)}{wp_1(y) + (1-w)p_2(y)}p_2(\lambda|y),
\end{aligned}
$$

where $p_1(\lambda|y)$ and $p_2(\lambda|y)$ are the posteriors induced by each of the mixture components (obtained in **a**).

Therefore, the posterior weights are

$$
\begin{aligned}
w^* &= \frac{wp_1(y)}{wp_1(y) + (1-w)p_2(y)}, \\
1 - w^* &= \frac{(1-w)p_2(y)}{wp_1(y) + (1-w)p_2(y)},
\end{aligned}
$$

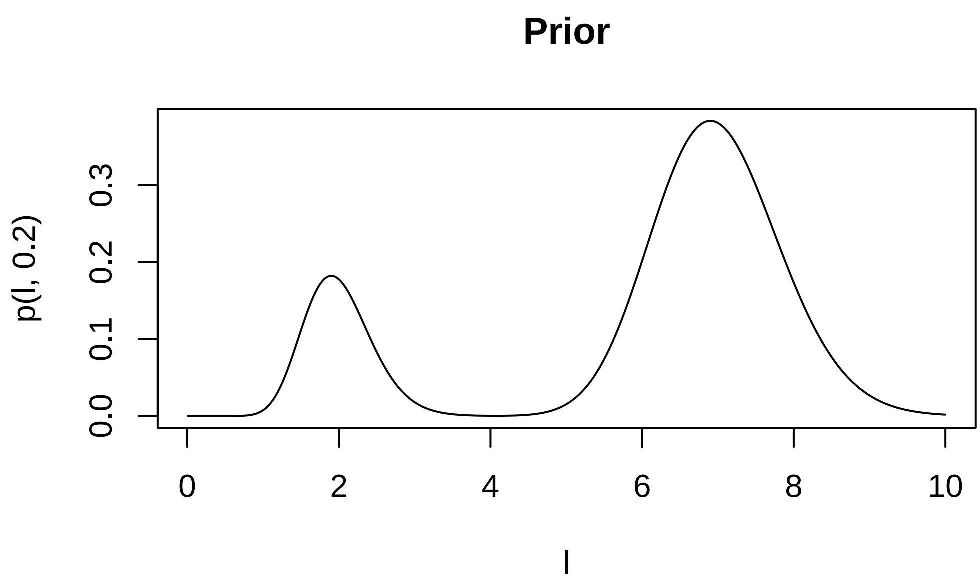
which are the prior weights adjusted by the prior predictive densities of the observed data.

c

Set $\alpha_1 = 20$, $\alpha_2 = 70$, $\beta_1 = \beta_2 = 10$, $w = 0.2$ and suppose we observe $y = 3$. Plot the prior and the posterior densities.

```
p <- function(l, w) {
  w*dgamma(l, 20, 10) + (1-w)*dgamma(l, 70, 10)
}
```

```r
l <- seq(from = 0.01, to = 10, by = 0.01)
plot(l, p(l, 0.2), type = "l", main = "Prior")
```



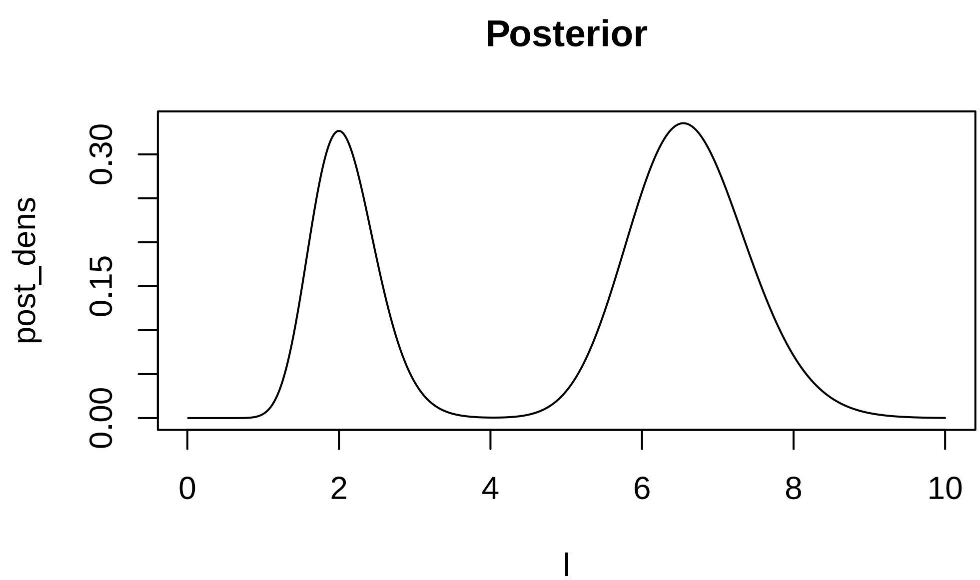

```r
wpost <- function(y, w, a1, b1, a2, b2) {
  num <- w*dnbinom(y, mu = a1/b1, size = b1)
  num/(num + (1-w)*dnbinom(y, mu = a2/b2, size = b2))
}
```

```r
y <- 3
a1 <- 20
b1 <- 10
a2 <- 70
b2 <- 10
w <- 0.2

wstar <- wpost(y, w, a1, b1, a2, b2)

posterior <- function(l) {
  wstar*dgamma(l, shape = a1 + y, rate = 1 + b1) +
  (1-wstar)*dgamma(l, shape = a2 + y, rate = 1 + b2)
```

```
}

integrate(posterior, 0, Inf)
```

```
1 with absolute error < 4.4e-06
```

```
post_dens <- posterior(l)
plot(l, post_dens, type = "l", main = "Posterior")
```



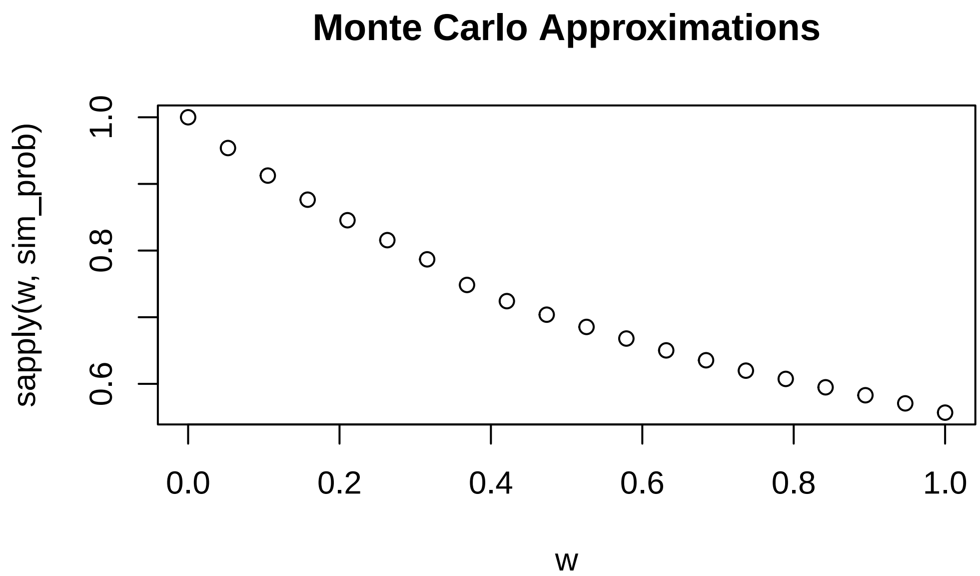
d

Using the parameters from part c, plot the Monte Carlo approximation of $\Pr(\lambda > 2|y)$ as a function of w, using 10,000 samples for each estimation.

```
sim_prob <- function(w, seed = 42) {
  set.seed(seed)
  wstar <- wpost(y, w, a1, b1, a2, b2)
  lambdas <- rep(0, times = 10000)
  for (i in 1:10000) {
    u <- rbinom(n = 1, size = 1, prob = wstar)
```

```r
    if (u == 1) {
      lambdas[[i]] <- rgamma(n = 1, shape = y + a1, rate = 1 + b1)
    } else {
      lambdas[[i]] <- rgamma(n = 1, shape = y + a2, rate = 1 + b2)
    }
  }

  return(mean(lambdas > 2))
}
```

```r
w <- seq(from = 0, to = 1, length.out = 20)
plot(w, sapply(w, sim_prob), main = "Monte Carlo Approximations")
```



2 Normal Model

Let $y_1, \ldots, y_n | \theta, \sigma^2 \sim_{iid} \mathcal{N}(\theta, \sigma^2)$. Consider a prior distribution for (θ, σ^2) of the form $p(\theta, \sigma^2) = p(\theta|\sigma^2)p(\sigma^2)$ with $\theta|\sigma^2 \sim \mathcal{N}(\mu_0, \tau_0^2)$. Obtain the full conditional distribution of θ, $p(\theta|\sigma^2, y_1, \ldots, y_n)$.

Likelihood:

$$
\begin{aligned}
p(y_1, \dots, y_n | \theta, \sigma^2) &= \prod_{i=1}^{n} p(y_i | \theta, \sigma^2) \\
&= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \theta)^2 \right)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{i=1}^{n} (y_i - \theta)^2 &= \sum_{i=1}^{n} (y_i - \bar{y} + \bar{y} - \theta)^2 \\
&= (n-1)s^2 + n(\bar{y} - \theta)^2
\end{aligned}
$$

Prior (proportionality as a function of θ):

$$
\begin{aligned}
p(\theta, \sigma^2) &\propto p(\theta | \sigma^2) \\
&\propto \exp\left(-\frac{1}{2\tau_0^2} (\theta - \mu_0)^2 \right)
\end{aligned}
$$

Full conditional:

$$
\begin{aligned}
p(\theta | \sigma^2, y_1, \dots, y_n) &\propto p(y_1, \dots, y_n | \theta, \sigma^2) p(\theta | \sigma^2) \\
&\propto \exp\left(-\frac{1}{2\sigma^2} n(\bar{y} - \theta)^2 \right) \times \exp\left(-\frac{1}{2\tau_o^2} (\theta - \mu_0)^2 \right) \\
&\propto \exp\left(-\frac{1}{2} \left(\frac{n(\bar{y} - \theta)^2}{\sigma^2} + \frac{(\theta - \mu_0)^2}{\tau_0^2} \right) \right) \\
&\propto \exp\left(-\frac{1}{2} \left[\theta^2 (n/\sigma^2 + 1/\tau_0^2) - 2\theta (n\bar{y}/\sigma^2 + \mu_0/\tau_0^2) \right] \right) \\
&\propto \exp\left(-\frac{1}{2} \left[\frac{n}{\sigma^2} + \frac{1}{\tau_0^2} \right] \left(\theta^2 - 2\theta \frac{\bar{y}n/\sigma^2 + \mu_0/\tau_0^2}{n/\sigma^2 + 1/\tau_0^2} \right) \right) \\
&\propto \exp\left(-\frac{1}{2} \left[\frac{n}{\sigma^2} + \frac{1}{\tau_0^2} \right] \left(\theta - \frac{\bar{y}n/\sigma^2 + \mu_0/\tau_0^2}{n/\sigma^2 + 1/\tau_0^2} \right)^2 \right),
\end{aligned}
$$

so

$$
\theta | \sigma^2, y_1, \dots, y_n \sim \mathcal{N}(\mu_n, \tau_n^2)
$$

with

$$
\begin{aligned}
\tau_n^2 &= \frac{1}{n/\sigma^2 + 1/\tau_0^2} \\
\mu_n &= \tau_n^2 \left(\frac{n}{\sigma^2} \bar{y} + \frac{1}{\tau_0^2} \mu_0 \right).
\end{aligned}
$$