Chapter 2 Solutions

2.1 Suppose that we want to develop an informative prior distribution for the probability of observing heads when we flip a coin. Suppose that we think that the most likely probability of heads is 0.5 and that 0.75 would be "extreme." Find the parameters of a beta density so that the median is approximately 0.5 and the 0.9 quantile is 0.75.

The following R code can be used to solve this problem. q is a vector of length 2 of quantiles, p is vector of length 2 of respective probabilities, *init* is a vector of length 2 of starting parameter values.

For this problem, $\alpha = \beta = 0.374$.

```
parametersolver = function(qu,p,init) {
  qu <- qu
  p <- p

  betaoptim = function(param) {
    q1 <- qu[1]
    q2 <- qu[2]
    p1 <- p[1]
    p2 <- p[2]
    (pbeta(q1,param[1],param[2])-p1)^2 + (pbeta(q2,param[1],param[2])-p2)^2
  }

  r = optim(init,betaoptim)

  v = unlist(r)
  t = c(v[1],v[2])
  print(t)
}

parametersolver(c(.5,.9),c(.5,.75),c(1,1))
```

2.2 Suppose that we are going to flip a coin 20 times.

(a) Using a beta distribution, write down a prior density that describes your uncertainty about the probability of "heads."

Using the prior from problem 2.1, we might use a Beta(0.374, 0.374). This distribution has median is 0.50, the 0.025 quantile is .00022, and the 0.975 quantile is 0.9998. This means there is 50% chance that the probability of heads is below 0.5, and it is very unlikely (a 5% chance) that the probability of getting heads will fall outside the interval (0.00022, 0.99998). The pdf of this distribution is

$$p(\pi) = \frac{\Gamma(0.748)}{\Gamma(0.374)\Gamma(0.374)}\pi^{0.374-1}(1-\pi)^{0.374-1} \qquad 0 \le \pi \le 1$$

(b) Flip a coin 20 times and record the outcomes. Write down the likelihood function for the observed data.

Let Y = number of heads in 20 flips of a coin. Then Y $\sim$ Binomial(20, π). My 20 flips of penny had 11 heads, 9 tails. So the likelihood is given by

$$Lik(\pi) = \binom{20}{11}\pi^{11}(1-\pi)^9$$

(c) Calculate the maximum likelihood estimate for the probability of "heads" and a 95% confidence interval.

$$\begin{aligned}
\log(Lik(\pi)) &= l(\pi) = log\binom{20}{11} + 11log(\pi) + 9log(1-\pi) \\
l'(\pi) &= \frac{11}{\pi} - \frac{9}{1-\pi} = 0 \\
\Rightarrow \hat{\pi} &= \frac{11}{20} \\
s.e.(\hat{\pi}) &= \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = \sqrt{\frac{.55(.45)}{20}} = 0.111
\end{aligned}$$

A 95% confidence interval is given by $\hat{\pi} \pm z_{0.975} s.e.(\hat{\pi}) = (0.332, 0.768)$

(d) Calculate the posterior distribution for the probability of "heads" and a 95% credible interval.

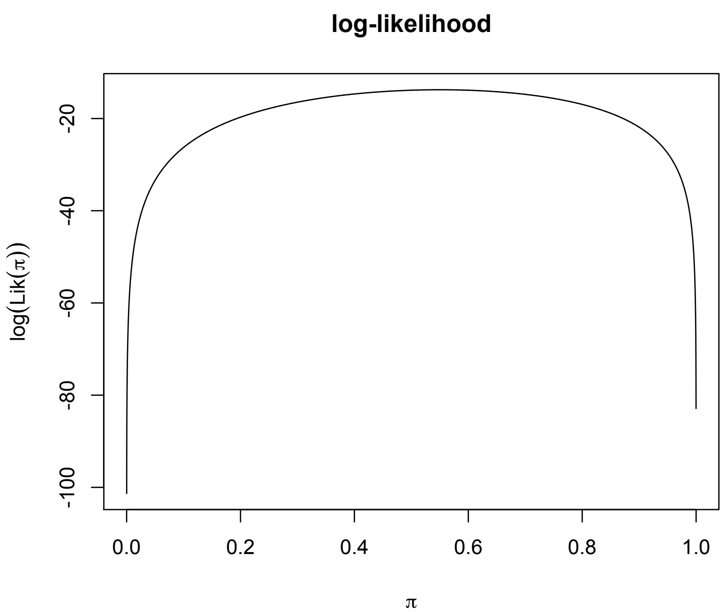
$$\begin{aligned}
p(\pi \,|\, y = 11) &\propto \pi^{11}(1-\pi)^9\pi^{0.374-1}(1-\pi)^{0.374-1} \\
&\propto Beta(11.374, 9.374)
\end{aligned}$$

Using R to find the 95% credible interval yields (0.337, 0.751).

```
qbeta(.025,11.374,9.374)
qbeta(.975,11.374,9.374)
```

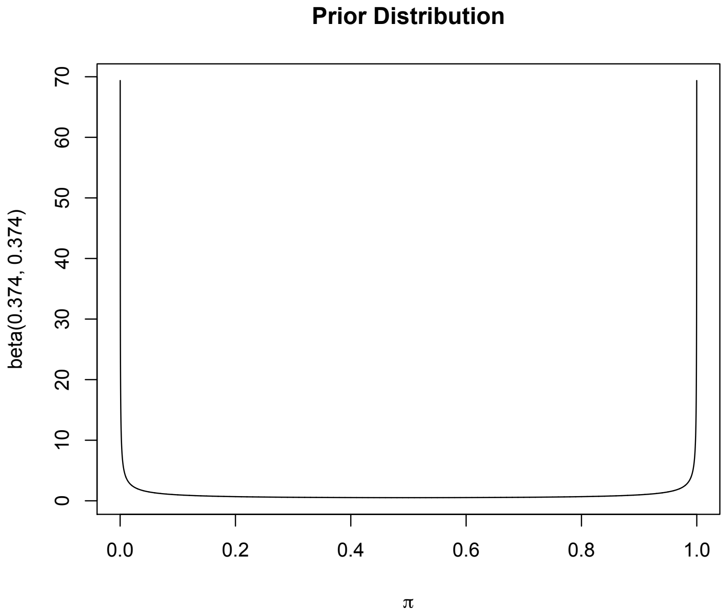
(e) Plot the log-likelihood function.

```
pi <- seq(0,1, by=.0001)
l <- 11*log(pi) + 9*log(1-pi)
plot(pi,l, type='l', main='log-likelihood', xlab=expression(pi),
     ylab=expression(Log-lik(pi)))
```



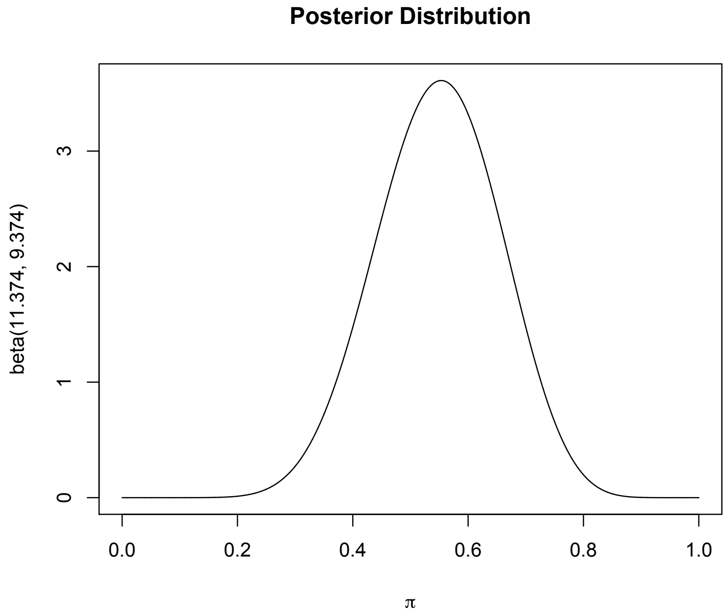
(f) Plot the prior density.

```
alpha <- .374
beta <- .374
beta <- (1/beta(alpha,beta)) * pi^(alpha-1) * (1-pi)^(beta-1)
plot(pi, beta, type='l', main='Prior Distribution',
     xlab=expression(pi), ylab='beta(0.374, 0.374)')
```



(g) Plot the posterior density.

```
alpha <- 11.374
beta <- 9.374
beta <- (1/beta(alpha,beta)) * pi^(alpha-1) * (1-pi)^(beta-1)
plot(pi, beta, type='l', main='Posterior Distribution',
     xlab=expression(pi), ylab='beta(11.374, 9.374)')
```



(h) Calculate the Bayes' factor comparing a uniform prior density to your informative prior density.

Let M_1 be the model with a uniform prior and M_2 be the model with a Beta(0.374, 0.374) prior. Then

$$
\begin{aligned}
m_1(11 \mid M_1) &= \int_0^1 \binom{20}{11} \pi^{11}(1-\pi)^9 d\pi \\
&= \binom{20}{11} \frac{\Gamma(12)\Gamma(10)}{\Gamma(22)} \int_0^1 \frac{\Gamma(22)}{\Gamma(12)\Gamma(10)} \pi^{12-1}(1-\pi)^{10-1} d\pi \\
&= 1/21 \\
&= 0.04761905
\end{aligned}
$$

$$
\begin{aligned}
m_2(11 \mid M_2) &= \int_0^1 \binom{20}{11} \pi^{11}(1-\pi)^9 \pi^{0.374-1}(1-\pi)^{0.374-1} d\pi \\
&= \binom{20}{11} \frac{\Gamma(11.374)\Gamma(9.374)}{\Gamma(20.748)} \\
&= 0.1175762
\end{aligned}
$$

So the Bayes' factor in favor of M_1 is $\frac{0.04761905}{0.1175762} = 0.405$.

2.3 Consider again the fluid breakdown times introduced in Sect. 2.5. Two models were proposed for these data. The first incorporated a normal likelihood function and a noninformative prior distribution; the second a normal likelihood function and a conjugate inverse-gamma/normal prior distribution. Now suppose that the properties of the manufacturing process were controlled when these samples of lubricant were produced so that it is known that the true mean of the sample values must lie between 6.0 and 7.4 (on the original measurement scale). No further information is available concerning the value of the variance parameter σ^2.

(a) Assume that the joint prior distribution for (μ, σ^2) is proportional to $1/\sigma^2$ whenever $\mu \in (\log(6.0), \log(7.4))$, and is 0 otherwise. Find an expression for a function that is proportional to the joint posterior distribution.

$$p(\mu, \sigma^2 \,|\, \vec{y}) \propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right] I[\mu \in (\log(6.0), \log(7.4))]$$

(b) Find a function that is proportional to the marginal posterior distribution of μ.

$$p(\mu \,|\, \vec{y}) \propto \left[1 + \frac{(\mu-\bar{y})^2}{(n-1)s^2/n}\right]^{-n/2} I[\mu \in (\log(6.0), \log(7.4)]$$

(c) Find a function that is proportional to the marginal posterior distribution of σ^2.

$$p(\sigma^2 \,|\, \vec{y}) \propto \int_{\log(6.0)}^{\log(7.4)} (\sigma^2)^{-n/2-1} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right] d\mu$$

2.4 Show that the beta distribution is the conjugate prior distribution for the binomial likelihood.

Suppose that we have data $X \sim Binomial(n, \pi)$, and that the prior distribution for π is $Beta(\alpha, \beta)$. Then the posterior distribution of π is given as

$$\begin{aligned} p(\pi \mid x, \alpha, \beta) &\propto \pi^x (1-\pi)^{n-x} \pi^{\alpha-1} (1-\pi)^{\beta-1} \\ &\propto \pi^{x+\alpha-1} (1-\pi)^{n-x+\beta-1}, \end{aligned}$$

which is $Beta(x + \alpha, n - x + \beta)$.

2.5 Show that the gamma distribution is the conjugate prior distribution for the mean of a Poisson likelihood.

Suppose that we have data $X_i \sim Poisson(\lambda)$, and that the prior distribution for λ is $Gamma(\alpha, \beta)$. Then the posterior distribution of π is given as

$$\begin{aligned} p(\lambda \mid \vec{x}, \alpha, \beta) &\propto \lambda^{\sum_{i=1}^n x_i} \exp(-n\lambda) \lambda^{\alpha-1} \exp(-\beta\lambda) \\ &\propto \lambda^{\alpha + \sum_{i=1}^n x_i - 1} \exp(-(\beta + n)\lambda), \end{aligned}$$

which is $Gamma(\alpha + \sum_{i=1}^n x_i, n + \beta)$.

2.6 Show that the gamma distribution is the conjugate prior distribution for the exponential likelihood.

Suppose that we have data $X_i \sim Exponential(\lambda)$, and that the prior distribution for λ is $Gamma(\alpha, \beta)$. Then the posterior distribution of π is given as

$$\begin{aligned} p(\lambda \mid \vec{x}, \alpha, \beta) &\propto \lambda^n \exp(-\lambda \sum_{i=1}^n x_i) \lambda^{\alpha-1} \exp(-\beta\lambda) \\ &\propto \lambda^{n+\alpha-1} \exp(-(\beta + \sum_{i=1}^n x_i)\lambda), \end{aligned}$$

which is $Gamma(\alpha + n, \beta + \sum_{i=1}^n x_i)$.

2.7 Derive the mean and variance for the lognormal distribution.

$$\begin{aligned}
E[X] &= \int_0^\infty x \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(\log(x)-\mu)^2\right] dx \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}y^2)\exp(y+\mu)dy \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(y-\sigma^2)^2)\exp(\mu+\frac{\sigma^2}{2})dy \\
&= \exp(\mu+\frac{\sigma^2}{2}).
\end{aligned}$$

$$\begin{aligned}
E[X^2] &= \int_0^\infty x^2 \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(\log(x)-\mu)^2\right] dx \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}y^2)\exp(2y+2\mu)dy \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(y-2\sigma^2)^2)\exp(2\mu+2\sigma^2)dy \\
&= \exp(2\mu+2\sigma^2).
\end{aligned}$$

The mean of the lognormal distribution is $E[X] = \exp(\mu + \frac{\sigma^2}{2})$, and the variance is $Var[X] = E[X^2] - (E[X]^2) = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2)$.

2.8 Suppose we are using an Exponential(λ) distribution to model the lifetimes of n items.

(a) Find the maximum likelihood estimator of λ.

The log-likelihood function for λ is $\log(L(\lambda \,|\, \vec{t})) = n\log(\lambda) - \lambda\sum_{i=1}^n t_i$. Taking derivatives,

$$\frac{d}{d\lambda}\log(L(\lambda \,|\, \vec{t})) = \frac{n}{\lambda} - \sum_{i=1}^n t_i.$$

Setting the derivative equal to zero and solving for λ gives $\frac{n}{\lambda} = \sum_{i=1}^n t_i$, or $\hat{\lambda} = \frac{n}{\sum_{i=1}^n t_i}$.

(b) Assume n is large and find the standard error of $\hat{\lambda}$.

The second derivative of the log-likelihood is

$$\frac{d^2}{d\lambda^2} \log(L(\lambda \,|\, \vec{t})) = -\frac{n}{\lambda^2}.$$

The observed Fisher information is $\frac{\hat{\lambda}}{\sqrt{n}} = \frac{\sqrt{n}}{\sum_{i=1}^n t_i}$.

(c) Suppose that we observed $n = 50$ items and that $\sum_{i=1}^{50} t_i = 25$. Find a 90% confidence interval for λ.

A 90% confidence interval for λ is

$$(\hat{\lambda} - 1.645 se(\hat{\lambda}), \hat{\lambda} + 1.645 se(\hat{\lambda})),$$

which is $(50/25 - 1.645\sqrt{(}50)/25, 50/25 + 1.645\sqrt{(}50)/25) = (2 - 0.465, 2 + 0.465) = (1.535, 2.464)$.

(d) Suppose that $\lambda \sim Gamma(1, 2)$. Find the posterior distribution for λ.

From Exercise 2.6, the posterior distribution is $Gamma(\alpha + n, \beta + \sum_{i=1}^n x_i) = Gamma(1 + 50, 2 + 25) = Gamma(51, 27)$.

(e) Suppose that we observed $n = 50$ items and that $\sum_{i=1}^{50} t_i = 25$. What is the posterior probability that λ falls in the 90% confidence interval found in (c)?

$$\int_{1.535}^{2.464} \frac{27^{51}}{\Gamma(51)} \exp(-27x) x^{50} dx = 0.979.$$