Solutions to selected exercises in problem set 12

Exercise 9.29

a) In this exercise we consider a random sample of iid Bernouilli(p) random variables. A $1-\alpha$ credible set A for the parameter p is defined by

$$P(p \in A|\mathbf{x}) = \int_A \pi(p|\mathbf{x})dp = 1-\alpha.$$

We shall find an expression for A, but first we need to derive the posterior distribution $\pi(p|\mathbf{x})$. The prior we will use is given in the exercise: $p \sim \text{beta}(a,b) = \pi(p)$.

The posterior is defined by

$$\pi(p|\mathbf{x}) = \frac{f(\mathbf{x}|p)\pi(p)}{\int_0^1 f(\mathbf{x}|p)\pi(p)dp}.$$

We have that $f(\mathbf{x}|p)\pi(p) = \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i} = p^{\sum x_i}(1-p)^{n-\sum x_i}$, and for simplicity we write $y = \sum x_i$. Furthermore, $\pi(p) = \frac{1}{B(a,b)}p^{(a-1)}(1-p)^{(b-1)}$. The expression for the posterior is then

$$\begin{aligned}
\pi(p|\mathbf{x}) &= \frac{p^y(1-p)^{n-y}\frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}}{\int_0^1 p^y(1-p)^{n-y}\frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}dp} \\
&= \frac{p^{y+a-1}(1-p)^{n-y+b-1}\frac{1}{B(a,b)}}{\int_0^1 p^{y+a-1}(1-p)^{n-y+b-1}\frac{1}{B(a,b)}dp} \\
&= \frac{p^{y+a-1}(1-p)^{n-y+b-1}}{\int_0^1 p^{y+a-1}(1-p)^{n-y+b-1}dp}.
\end{aligned}$$

To solve the integral, recall the definition of the Beta-function

$$B(u,v) = \int_0^1 x^{u-1}(1-x^{v-1})dx.$$

If we set $u = y+a$ and $v = n-y+b$ we get the same form, and thereby

$$\pi(p|\mathbf{x}) = \frac{1}{B(y+a, n-y+b)}p^{y+a-1}(1-p)^{n-y+b-1},$$

which we recognize as the pdf of a beta-distributed variable with parameters $(y+a, n-y+b)$. So the prior distribution is a beta-distribution.

Note: In the exercise we were told to use the *conjugate beta(a,b) prior*. A conjugate prior means that the the posterior will be the same type of distribution, as we observed here. However, we needed to write out the definition of the posterior in order to find the correct parameters of the posterior distribution.

The beta-distribution is either strictly decreasing, increasing, U-shaped or unimodal. Therefore we can use an interval as our credible set. If we let $\beta_{\alpha/2}$ and $\beta_{1-\alpha/2}$ be critical values of the beta-distribution with parameters $(y+a, n-y+b)$, then $P(\beta_{1-\alpha/2} < p < \beta_{\alpha/2}) = 1-\alpha$. Formally, we write the credible set as

$$A = \{p : \beta_{1-\alpha/2} < p < \beta_{\alpha/2}\}.$$

Exercise 9.27

a) We have a random sample from an exponential(λ) pdf, and λ has a conjugate IG(a, b) prior (inverted gamma). We are going to *show how to find* a $1-\alpha$ HPD credible set for λ. HDP stands for Highest Posterior Density. In other words, out of all the possible $1-\alpha$ credible intervals, we want to find the shortest one. We need corollary 9.3.10 to solve this exercise. The corollary states that if the posterior density $\pi(\lambda|\mathbf{x})$ is unimodal then the shortest credible interval for λ is

$$\{\lambda : \pi(\lambda|\mathbf{x}) \geq k\} \text{ where } k \text{ is such that } \int_{\{\lambda:\pi(\lambda|\mathbf{x})\geq k\}} \pi(\lambda|\mathbf{x})d\lambda = 1-\alpha$$

We start by finding the posterior. Note (again) that we have a conjugate prior so the posterior will also be an inverted gamma distribution. We need to find the parameters of this distribution. The posterior is defined by

$$\pi(\lambda|\mathbf{x}) = \frac{f(\mathbf{x}|\lambda)\pi(\lambda)}{\int_0^\infty f(\mathbf{x}|\lambda)\pi(\lambda)d\lambda} = \frac{f(\mathbf{x}|\lambda)\pi(\lambda)}{m(\mathbf{x})}.$$

We only need to work with this expression until we recognize the parameters of the posterior. Consider the following

$$f(\mathbf{x}|\lambda) = \prod_{i=1}^n f(x_i|\lambda) = \prod_{i=1}^n \frac{1}{\lambda}e^{-x_i/\lambda} = \left(\frac{1}{\lambda}\right)^n e^{-\frac{1}{\lambda}\sum_{i=1}^n x_i}$$

For simplicity we write $y = \sum_{i=1}^n x_i$. Then

$$\begin{aligned}
f(\mathbf{x}|\lambda)\pi(\lambda) &= \left(\frac{1}{\lambda}\right)^n e^{-\frac{1}{\lambda}y}\frac{1}{\Gamma(a)b^a}\left(\frac{1}{\lambda}\right)^{a+1} e^{-\frac{1}{\lambda}\frac{1}{b}} \\
&= \left(\frac{1}{\lambda}\right)^{n+a+1} e^{-\frac{1}{\lambda}(y+\frac{1}{b})}\frac{1}{\Gamma(a)b^a}
\end{aligned}$$

Note that we already see something similar to the form of a inverse gamma distribution, and our best guess would be that the parameters of the posterior are $n+a$ and $1/(y+\frac{1}{b})$. To find the posterior we divide by the function $m(\mathbf{x})$. Note that this function can contain expressions with y, a and b, but not λ! Therefore, the only thing that will change in the above expression when we divide by $m(\mathbf{x})$, is a constant with respect to λ. We can therefore already here conclude that our guess for the parameters of the posterior distribution is correct.

For completeness, I write out the posterior to show that the guess is in fact correct. The term $\frac{1}{\Gamma(a)b^a}$ will be canceled when we divide by $m(\mathbf{x})$. Then it only remains to calculate the integral

$$\int_0^\infty \left(\frac{1}{\lambda}\right)^{n+a+1} e^{-\frac{1}{\lambda}(y+\frac{1}{b})}d\lambda.$$

Looks difficult, but if our guess is correct is should be equal to $\Gamma(n+a)(1/(y+1/b))^{n+a}$, so that the posterior will be the pdf of a gamma distribution with parameters $n+a$ and $1/(y+\frac{1}{b})$.

The definition of the gamma function is

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt.$$

In our integral, make the substitution $t = \frac{1}{\lambda}(y + \frac{1}{b})$. Then $\frac{dt}{d\lambda} = -\frac{1}{\lambda^2}(y + \frac{1}{b}) = -t^2/(y + \frac{1}{b})$. Thus,

$$\begin{aligned}
\int_0^\infty \left(\frac{1}{\lambda}\right)^{n+a+1} e^{-\frac{1}{\lambda}(y+\frac{1}{b})}d\lambda &= -\int_\infty^0 \left(\frac{t}{y+\frac{1}{b}}\right)^{n+a+1} e^{-t}\left(\frac{y+\frac{1}{b}}{t^2}\right)dt \\
&= \left(\frac{1}{y+\frac{1}{b}}\right)^{n+a} \int_0^\infty t^{n+a+1}e^{-t}dt \\
&= \left(\frac{1}{y+\frac{1}{b}}\right)^{n+a} \Gamma(n+a).
\end{aligned}$$

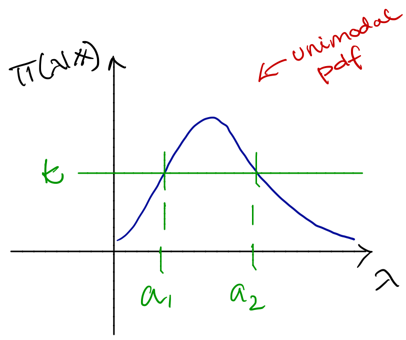
Now that we have the posterior distribution we can take another look at the corollary. First we need to know whether the posterior pdf is unimodal. The answer is yes, check wikipedia article on inverse gamma to see the shape of the pdf (you do not need to know the shape of all kinds of distributions for an exam). By corollary 9.3.10 we then know that the HPD credible set is given by

$$\{\lambda : \pi(\lambda|\mathbf{x}) \geq k\}.$$

How do we go from here to an interval for λ? See the figure below. We see that the HDP credible interval is $a_1 \leq \lambda \leq a_2$, where a_1 and a_2 solves the system of equations

$$\begin{aligned}
&\pi(a_1|y) = \pi(a_2|y) \rightarrow \left(\frac{1}{a_1}\right)^{n+a+1} e^{-\frac{1}{a_1}(y+\frac{1}{b})} = \left(\frac{1}{a_2}\right)^{n+a+1} e^{-\frac{1}{a_2}(y+\frac{1}{b})} \\
&P(a_1 < \lambda < a_2) = 1 - \alpha
\end{aligned}$$

We would find the solution by solving this set of equations. So we have now *showed how to find* a $1-\alpha$ HPD credible set for λ.



$\pi(\lambda|x) \geq k$ equivalent to $a_1 \leq \lambda \leq a_2$.

$k = \pi(a_1|x) = \pi(a_2|x)$