Homework 2 Solution

Math 315, Fall 2019

1.

Bob claims to have ESP (extrasensory perception). To test this claim, you propose the following experiment. You will select one from four large cards with different geometric figures and Bob will try to identify it. Let θ denote the probability that Bob is correct in identifying the figure for a single card. You believe that Bob has no ESP ability ($\theta = .25$), but there is a small chance that θ is either larger or smaller than .25. After some thought, you place the following prior distribution on θ:

θ	0	.125	.25	.375	.5	.625	.75	.875	1
$\pi(\theta)$	.001	.001	.95	.008	.008	.008	.008	.008	.008

Suppose that the experiment is repeated ten times and Bob is correct six times and incorrect four times.

(a) Derive the posterior distribution of θ and display it as a table.

I derived this in R, but you should also understand how to do this by hand.

```
n <- 10
y <- 6
theta <- c(0, .125, .25, .375, .5, .625, .75, .875, 1)
prior <- c(.001, .001, .95, .008, .008, .008, .008, .008, .008)
likelihood <- dbinom(y, size = n, prob = theta)
unstd.posterior  <- prior * likelihood
posterior <- prior * likelihood / sum(unstd.posterior)
```

I don't expect ridiculious precision to be reported, so here is the posterior to four decimal places.

theta	prior	likelihood	posterior
0.000	0.001	0.0000	0.0000
0.125	0.001	0.0005	0.0000
0.250	0.950	0.0162	0.7305
0.375	0.008	0.0891	0.0338
0.500	0.008	0.2051	0.0778
0.625	0.008	0.2475	0.0939
0.750	0.008	0.1460	0.0554
0.875	0.008	0.0230	0.0087
1.000	0.008	0.0000	0.0000

(b) What is your posterior probability that Bob has no ability (and is thus simply guessing)?

If Bob has no ability, then Bob will simply be guessing, making $\theta = 0.25$. The posterior probability, $P(\theta = 0.25|Y = 6)$, is approximately 0.73.

Note: I would also accept $P(\theta \leq 0.25|Y = 6)$ as a reasonable answer here, since the implied alternative hypothesis is then $\theta > 0.25$. In fact, this type of hypothesis is quite common in the Bayesian setting.

2.

A study reported on the long-term effects of exposure to low levels of lead in childhood. Researchers analyzed children's shed primary teeth for lead content. Of the children whose teeth had a lead content of more than 22.22 parts per million (ppm), 22 eventually graduated from high school and 7 did not. Suppose your prior density for θ, the proportion of all such children who will graduate from high school is Beta(1, 1).

(a) Derive the posterior distribution of θ.

The model here is

$$Y|\theta \sim \text{Binomial}(n, \theta)$$
$$\theta \sim \text{Beta}(1, 1)$$

We derived the posterior distribution in class when the prior is Beta and the likelihood is Binomial, so I will omit the full details here. You should have found the posterior to be Beta(23, 8), which you can see based on the below proportionality setting $Y = 22$ and $n = 29$:

$$p(\theta|Y) \propto \theta^Y (1-\theta)^{n-Y}$$

(b) Calculate the MAP estimate for θ.

(c) Use the R function `qbeta()` to find a 93% equal-tailed credible interval for θ.

```
qbeta(c(0.035, 0.965), 23, 8)
```

```
## [1] 0.5907769 0.8692106
```

There is approximately a 93% chance that the proportion of all children who were exposed to low-levels of lead in childhood gradutaing from high school is between 0.59 and 0.87 (given the data and the prior).

(d) Calculate a 93% HPDI.

```
library(HDInterval)
draws <- rbeta(10000, 23, 8)
hdi(draws, credMass = 0.93)
```

```
##     lower     upper
## 0.6011237 0.8728442
## attr(,"credMass")
## [1] 0.93
```

Notice that this interval is narrower than the equal-tailed credible interval.

(e) Use the function `pbeta()` to find the probability that θ exceeds 0.6.

```
1 - pbeta(0.6, 23, 8)
```

```
## [1] 0.9564759
```

3.

The Weibull distribution is often used as a model for survival times in biomedical, demographic, and engineering analyses. A random variable Y has a Weibull distribution if has PDF

$$f(y|\alpha,\lambda) = \lambda\alpha y^{\alpha-1}\exp(-\lambda y^{\alpha}),\ \text{where } y > 0,\ \alpha > 0,\ \lambda > 0.$$

For this problem, assume that α is known, but λ is unknown.

Assume $Y_1, Y_2, \ldots, Y_n$ are i.i.d. Weibull random variables with known α and that λ has prior $\lambda \sim \text{Gamma}(a, b)$. Derive the resulting posterior distribution. If it is a member of a named family of distributions, be sure to specify this, along with its parameter values.

The model here is

$$\begin{aligned} Y_{|}\theta &\overset{\text{iid}}{\sim} \text{Weibull}(\alpha,\lambda) \\ \lambda &\sim \text{Gamma}(a,b) \end{aligned}$$

So we have that

$$\begin{aligned} f(y_i|\alpha,\lambda) &= \lambda\alpha y^{\alpha-1}e^{-\lambda y^{\alpha}} \\ \pi(\lambda) &= \frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda} \end{aligned}$$

The posterior is

$$\begin{aligned} p(\lambda|y_1,\ldots,y_n) &\propto \left[\prod_{i=1}^{n}\lambda e^{-\lambda y^{\alpha}}\right]\lambda^{a-1}e^{-b\lambda} \\ &= \lambda^{(n+a)-1}e^{-\lambda\left[b+\sum y_i^{\alpha}\right]} \end{aligned}$$

which is the kernel of $\text{Gamma}(n+a,\, b+\sum y_i^{\alpha})$.