STAT 431 Homework 2 Solutions

Robert Garrett

Problem 1

Part a

θ represents the probability of one telephone call resulting in a sale. We consider this a success.

Part b

Y is the total number of unsuccessful calls, so find $E(Y)$ and add the 5 successful calls.

$E(Y) = \frac{m(1-\theta)}{\theta} = \frac{5*0.8}{0.2} = 20$

Expected total number of unsuccessful calls is 20. Adding in 5 successful calls, we get 25 expected total calls.

Part c

$f(y|\theta) \propto \theta^m (1-\theta)^y = \theta^5 (1-\theta)^{20}$

Part d

$p(\theta|y) \propto f(y|\theta)\pi(\theta) \propto \theta^5 (1-\theta)^{20} * 1$

Part e

Factoring the above posterior, we get:

$p(\theta|y) \propto \theta^5 (1-\theta)^{20} = \theta^{6-1}(1-\theta)^{21-1}$

So, our posterior distribution is $\theta|y \sim Beta(6, 21)$

Problem 2

Part a

Comparing to Part e of Problem 1, we see the only difference is this posterior has an additional $(1-\theta)^4$ factor. Since the previous uniform prior doesn't contribute any terms with θ included, we need only replace it with a prior with a kernel of $(1-\theta)^4$. The $Beta(\alpha = 1, \beta = 5)$ prior has this kernel, so choose this as $\pi(\theta)$:

$$
\begin{aligned}
p(\theta|y) &\propto f(y|\theta)\pi(\theta) \\
&\propto \theta^5(1-\theta)^{20}\theta^0(1-\theta)^4 \\
&= \theta^5(1-\theta)^{24} = \theta^{6-1}(1-\theta)^{25-1}
\end{aligned}
$$

This matches the desired posterior, so the analyst used a $Beta(1, 5)$ prior.

Part b

```
alpha = 6; beta = 25
(post.mean = alpha/(alpha+beta))
```

```
## [1] 0.1935484
```

```
(post.sd = sqrt(alpha*beta/( ((alpha+beta)^2)*(alpha+beta+1) )))
```

```
## [1] 0.06984076
```

So, our posterior mean is approximately 0.194 and our posterior standard deviation is approximately 0.070.

Part c

```
qbeta(c(0.025,0.975),alpha,beta)
```

```
## [1] 0.07713551 0.34721170
```

95% equal-tailed credible interval for θ is approximately $(0.077, 0.347)$.

Part d

```
pbeta(0.2,alpha,beta)
```

```
## [1] 0.5724876
```

So, our posterior probability that $\theta \leq 0.2$ is approximately 0.572.

Part e

Yesterday, the salesperson made 20 unsuccessful calls to reach the quota. We have used this data to obtain posterior estimates of θ. We now use this posterior to find the posterior predictive distribution:

$$
\begin{aligned}
f(y^*|y) &= \int_0^1 f(y^*|\theta)p(\theta|y)d\theta \\
&= \binom{y^*+4}{y^*}\frac{\Gamma(31)}{\Gamma(6)\Gamma(25)}\int_0^1 \theta^5(1-\theta)^{y^*}\theta^5(1-\theta)^{24}d\theta \\
&= \binom{y^*+4}{y^*}\frac{\Gamma(31)}{\Gamma(6)\Gamma(25)}\frac{\Gamma(11)\Gamma(25+y^*)}{\Gamma(36+y^*)}\int_0^1 \frac{\Gamma(36+y^*)}{\Gamma(11)\Gamma(25+y^*)}\theta^{10}(1-\theta)^{y^*+24}d\theta \\
f(y^*|y) &= \binom{y^*+4}{y^*}\frac{\Gamma(31)}{\Gamma(6)\Gamma(25)}\frac{\Gamma(11)\Gamma(25+y^*)}{\Gamma(36+y^*)} * 1, \quad y^* = 0,1,2,...
\end{aligned} \tag{1}
$$

Now, we sum over the posterior predictive density function for the desired values.

```
posterior_predictive_density = function(y_star){
  choose(y_star+4,y_star)*(beta(6,25)^-1)*beta(11,25+y_star)
}
sum(posterior_predictive_density(0:19))
```

```
## [1] 0.4810545
```

There is a posterior predictive probability of approximately 0.481 that the salesperson will need to make strictly fewer (less than 20) calls today in order to meet the quota.

Problem 3

Part a

First, find the posterior distribution:

$$\begin{aligned}p(\theta|y) &\propto f(y|\theta)\pi(\theta)\\&\propto \theta^y(1-\theta)^{n-y}\theta^{\alpha-1}(1-\theta)^{\beta-1}\\&= \theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1}\end{aligned}$$

This is the kernel of the Beta distribution, so we see $\theta|y \sim Beta(y+\alpha, n-y+\beta)$.

Now, we use our knowledge of the variance of the Beta distribution from the appendix of BSM to find the posterior standard deviation (the y terms in the denominator cancel out):

$$\sqrt{Var(\theta|y)} = \sqrt{\frac{(y+\alpha)(n-y+\beta)}{(\alpha+n+\beta)^2(\alpha+n+\beta+1)}}$$

Part b

Looking at the previous posterior standard deviation, we see that there are two y terms remaining. Since we are now treating the data as random, and Y can range from $0, 1, 2, ..., n$, we'll need to factor each Y term into $\hat{\theta} = Y/n$ in order to analyze the convergence. So, for $0 < \theta < 1$:

$$\begin{aligned}
lim_{n\to\infty}\frac{\sqrt{Var(\theta|y)}}{SE(\hat{\theta})} &= lim_{n\to\infty}\frac{\sqrt{\frac{(Y+\alpha)(n-Y+\beta)}{(\alpha+n+\beta)^2(\alpha+n+\beta+1)}}}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} && (2)\\
&= lim_{n\to\infty}\frac{\sqrt{\frac{n^2(Y/n+\alpha/n)(n/n-Y/n+\beta/n)}{n^3(\alpha/n+n/n+\beta/n)^2(\alpha/n+n/n+\beta/n+1/n)}}}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} && (3)\\
&= lim_{n\to\infty}\frac{\sqrt{\frac{(\hat{\theta}+\alpha/n)(1-\hat{\theta}+\beta/n)}{n(\alpha/n+1+\beta/n)^2(\alpha/n+1+\beta/n+1/n)}}}{\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} && (4)\\
&= lim_{n\to\infty}\frac{\sqrt{\frac{(\hat{\theta}+\alpha/n)(1-\hat{\theta}+\beta/n)}{(\alpha/n+1+\beta/n)^2(\alpha/n+1+\beta/n+1/n)}}}{\sqrt{\hat{\theta}(1-\hat{\theta})}} && (5)\\
&= \frac{\sqrt{\frac{(\theta+0)(1-\theta+0)}{(0+1+0)^2(0+1+0+0)}}}{\sqrt{\theta(1-\theta)}}, \quad w.p.\ 1 && (6)\\
&= \frac{\sqrt{\theta(1-\theta)}}{\sqrt{\theta(1-\theta)}}, \quad w.p.\ 1 && (7)\\
&= 1, \quad w.p.\ 1 && (8)
\end{aligned}$$

Line 2: factor n out of each term in the numerator. Line 3: $Y/n = \hat{\theta}$. Line 4: Cancel $\sqrt{\frac{1}{n}}$ from numerator and denominator. Line 5: Pass limit inside the function (function is continuous); apply SLLN to get θ without the hat; α, β are finite so $\alpha/n, \beta/n \to 0$.