

PS9–MATH392

Theodore Dounias

4/15/2018

10.1

a.

```
theta <- c(.4, .5, .6)
prior <- c(1/5, 3/5, 1/5)
lik <- theta^4 * (1-theta)
pxl <- prior*lik
post <- pxl/sum(pxl)

data <- data.frame(theta, prior, lik, pxl, post)

data %>%
  kable() %>%
  kable_styling()
```

theta	prior	lik	pxl	post
0.4	0.2	0.01536	0.003072	0.0954334
0.5	0.6	0.03125	0.018750	0.5824790
0.6	0.2	0.05184	0.010368	0.3220876

- b. Since posterior probabilities can only be interpreted comparatively, the posterior spikes at .5, and then declines by about half its value at .6, which could be interpreted as evidence against the .6 value.
- c. The prior expected value is $E(\theta) = \sum_1^3 \theta_i \times prior = 0.5$
The posterior expected value is $E(\theta|X) = \sum_1^3 \theta_i \times posterior = 0.5226654$

10.5

- a. I operate under the assumption that each individual's response is an iid Bernoulli random variable with a set probability p, so that the total number of positive responses is a Binom(1324, p). A good estimate for p is x/n, where x = 295.

```
n <- 1324
x <- 295
p <- x/n

prop.test(x, n, p = p, conf.level = .90)[6]

## $conf.int
## [1] 0.2045745 0.2421753
## attr(,"conf.level")
## [1] 0.9
```

- b. For binomial data, the posterior follows a beta distribution: Beta(a+x, b+n-x)

```
#Find beta parameters using function I made
par <- calc.param.beta(.3, .02)
```

```
bayes.ci <- c(qbeta(.05, par[1]+x, par[2]+n-x), qbeta(.95, par[1]+x, par[2]+n-x))
bayes.ci
```

```
## [1] 0.2284099 0.2612982
```

c. Since theta given the data follows a beta (R H Y M E S), I can use the cdf:

```
pbeta(.23, par[1]+x, par[2]+n-x)
```

```
## [1] 0.06930314
```

10.8

a. I apply the equations on p.317 of the textbook. I will find m1, s1, the parameters of the posterior normal:

```
xbar <- 538
n <- 60
m0 <- 600
s0 <- 25
s <- 116

calc.param.norm <- function(xbar, n, m0, s0, s){
  m1 <- (1/s0^2)/(s0^(-2) + n*s^(-2))*m0 + (n/s^2)/(s0^(-2) + n*s^(-2))*xbar
  s1 <- 1/(sqrt(s0^(-2) + n*s^(-2)))
  out <- c(m1, s1)
  out
}

m1 <- calc.param.norm(xbar, n, m0, s0, s)[1]
s1 <- calc.param.norm(xbar, n, m0, s0, s)[2]
```

b. The CI is as follows:

```
ci <- c(qnorm(.025, m1, s1), qnorm(.975, m1, s1))
ci
```

```
## [1] 529.1928 579.5520
```

c.

```
1-pnorm(600, m1, s1)
```

```
## [1] 0.0001914289
```

10.11

A. (a, b, c) I will use the previously defined function.

```
xbar <- 19
s <- 6
m0 <- 25

res11 <- function(n, s0){
  c <- calc.param.norm(xbar, n, m0, s0, s)
  m1 <- c[1]
  s1 <- c[2]
  pre <- 1/s1
  out <- c(m1, s1, pre)
```

```

  out
}

parameter <- c("Mean", "Standard Deviation", "Precision")

dt <- data.frame(parameter, res11(15, 5), res11(50, 5), res11(15, 10))

names(dt) <- c("Parameter", "A", "B", "C")

dt %>%
  kable() %>%
  kable_styling()

```

Parameter	A	B	C
Mean	19.5255474	19.1679627	19.1406250
Standard Deviation	1.4797909	0.8365671	1.5309311
Precision	0.6757712	1.1953614	0.6531973

Changes in standard deviation (Case A to Case C) have little to no effect; there is a very slight effect on the mean, but it is negligible if compared to the almost double standard deviation. Changes in sample size have a large effect, since Case B is the most precise with approximately half the standard deviation of the other two.

10.15

- From page 141, ex 6.4, the likelihood function for $\text{Unif}[0, \theta]$ is $L(\theta) = \left(\frac{1}{\theta}\right)^n$
- Here the prior is $\pi(\theta|\alpha, \beta)$. I then calculate:

$$P(\theta|X) = \frac{\alpha\beta^\alpha\theta^{-(\alpha+n+1)}}{\int_{\beta}^{+\infty} \alpha\beta^\alpha\theta^{-(\alpha+n+1)}d\theta} = \frac{\theta^{-(\alpha+n+1)}}{(\theta^{\alpha+n}(\alpha+n))^{-1}\Big|_{\beta}^{+\infty}} = (\alpha+n)\frac{\beta^{\alpha+n}}{\theta^{(\alpha+n)+1}}$$

Therefore $[\theta|X] \sim \text{Pareto}(\alpha+n, \beta)$, and the Pareto distribution family is conjugate to the uniform.

- The posterior will follow a $\text{Pareto}(\alpha+n, \beta)$. Inserting values for the parameters I have:

```
1 - ppareto(15, 4.3, 5)
```

```
## [1] 0.04348734
```

10.17

- From Chapter 6 $L(\theta) = \theta^n e^{-\theta} \sum x_i$
- Using similar calculations to the previous problem:

$$P(\theta|X) = \frac{\theta^n e^{-\theta} \sum x_i \frac{\lambda^r}{\Gamma(r)} e^{-\lambda\theta} \theta^{r-1}}{\int_0^\infty \theta^n e^{-\theta} \sum x_i \frac{\lambda^r}{\Gamma(r)} e^{-\lambda\theta} \theta^{r-1} d\theta} = \frac{\theta^{n+r-1} e^{-\theta(\sum x_i + \lambda)}}{\int_0^\infty \theta^{n+r-1} e^{-\theta(\sum x_i + \lambda)} d\theta} = \frac{(\sum x_i + \lambda)^{n+r}}{\Gamma(n+r)} e^{-(\sum x_i + \lambda)\theta} \theta^{n+r-1}$$

Therefore, $[\theta|X] \sim \text{Gamma}(n+r, \sum x_i + \lambda)$

- The posterior is $\text{Gamma}(14, 17)$
- Using R's quantile functions:

```
ci <- c(qgamma(.025, shape=14, rate=17), qgamma(.975, shape=14, rate=17))
ci
## [1] 0.4502312 1.3076703
```