Homework 3

Graham Gibson

Problem 1

Let

$$Y_1, Y_2, ..., Y_n \sim N(\mu, \sigma^2)$$

For now assume $\sigma = 15, \bar{y} = 113, n = 10$

We have

$$p(\mu) = N(\mu_0, \sigma_0^2) = N(100, 15)$$

We know from the slides that

$$E(\mu|y) = \frac{\frac{\mu_0}{15} + \frac{n\bar{y}}{15}}{\frac{n+1}{15}}$$

We can see that the σ terms drop out and we are left with

$$E(\mu|y) = \frac{\mu_0 + n\bar{y}}{n+1} = \frac{100 + 10 * 113}{11} = 111.818$$

Since we know that $\mu|y$ is normally distributed we can construct a 95% credible interval based on

$$\frac{\mu_0 + n\bar{y}}{n+1} - 1.96 * \sigma, \frac{\mu_0 + n\bar{y}}{n+1} + 1.96 * \sigma$$
$$\frac{100 + 10(113)}{11} - 1.96 * \sigma, \frac{\mu_0 + n\bar{y}}{n+1} + 1.96 * \sigma$$

where

$$\sigma^2 = \frac{1}{\frac{1}{15} + \frac{n}{15}} = \frac{15^2}{n+1} = V = \frac{15^2}{11} = 20.5$$

so our final credible interval is

```r
print ("Lower bound")
```

```
## [1] "Lower bound"
```

```r
print (signif((qnorm(.025,mean=111.818, sd = sqrt((15^2)/11))),5))
```

```
## [1] 102.95
```

```r
print ("Upper bound")
```

```
## [1] "Upper bound"
```

```r
print (signif((qnorm(.975,mean=111.818, sd = sqrt((15^2)/11))),5))
```

```
## [1] 120.68
```

Problem 2

We know by definition that bias is

$$E(\hat{\mu}|\mu^*) - \mu^*$$

Let's take a closer look at

$$E(\hat{\mu}|\mu^*)$$

$$= E(\frac{\mu_0 + n\bar{y}}{n+1}|\mu*) = \frac{\mu_0}{n+1} + \frac{n}{n+1}\mu^*$$

by

$$E(\bar{y}) = \mu^*$$

We can see that the Bayesian estimator is biased, unlike the frequentist estimator.

In particular, when $\mu^* = 112$ and $\mu_0 = 100$ we get the bias is

$$\frac{100}{11} + \frac{10}{11}112 - 112 = -1.09$$

The bias of the maximum likelihood estimate is given by $E(\bar{y}) - \mu^* = 0$ so the bayesian estimate has larger bias.

The variance of the bayes estimate is given by

$$Var(\frac{\mu_0 + n\bar{y}}{n+1}|\mu*) = \frac{1}{11^2}\frac{\sigma^2}{11} = 10\frac{225}{11^2} = 18.6$$

whereas the variance of the MLE estimate is given by

$$Var(\bar{y}) = \frac{225}{10} = 22.5$$

Putting this together we get that

$$MSE(Bayes) = 18.6 + 1.19 = 19.79$$

$$MSE(MLE) = 22.5$$

So the $Bias_{Bayes} > Bias_{Mle}$ but $Var_{Bayes} << Var_{Mle}$ so $MSE_{Bayes} < MSE_{MLE}$

```
y_bar <- 113
n <- 10
mu0 <- 100
sig2_0 <- 15^2
sig2_mu0 <- 15^2
var.y <- 13^2
nu0 <- 1
S <- 1000
nun <- nu0+n

PHI<- matrix(nrow=S,ncol=2)
PHI[1,]<- phi<- c(y_bar,1/var.y)
```

```r
# Gibbs sampling
set.seed (1)
for(s in 2:S){
# generate a new mu value from its full conditional
  sigma2.mun <- 1/(1/sig2_mu0 + n*phi[2])
  mun <- sigma2.mun*( mu0/ sig2_mu0 + n*y_bar*phi[2])
  phi[1] <- rnorm( 1 , mun, sqrt(sigma2.mun))

# generate a new 1/sigma^2 value from its full conditional
  sig2_n <-(1/nun)*(nu0*sig2_0 + (n-1)*var.y + n*(y_bar - phi[1])^2)
  phi[2] <- rgamma(1, nun/2 , nun*sig2_n/2)
  PHI[s,]<- phi
}
```

$\hat{\mu}$

```
## [1] 112
```

```
## [1] "Mu 95% Credible Interval"
```

```
##  2.5% 97.5%
##   102   120
```

```
## [1] "Sigma point estimate"
```

```
## [1] 13.4
```

```
## [1] 15.1
```

```
## [1] 14.6
```

```
## [1] 14.6
```

```
## [1] "Sigma 95% Credible Interval"
```

```
##  2.5% 97.5%
##  9.29 24.24
```

```
##     97.5%      2.5%
##  9.288552 24.237019
```