Final Exam

Miles Keppler

March 19, 2018

1

Selecting a beta distribution because $0 \leq \theta \leq 1$. The $\mathrm{E}[\theta] = \alpha/\beta = .03$ and $2\sqrt{\mathrm{var}(\theta)} = 2\sqrt{\alpha/\beta^2} = .07$.

$$\beta = \frac{\mathrm{E}[\theta]}{\mathrm{var}(\theta)} \qquad \alpha = \beta\,\mathrm{E}[\theta]$$
$$= \sqrt{\frac{.07}{2}} \qquad = .187 * .03$$
$$= .187 \qquad = .00561.$$

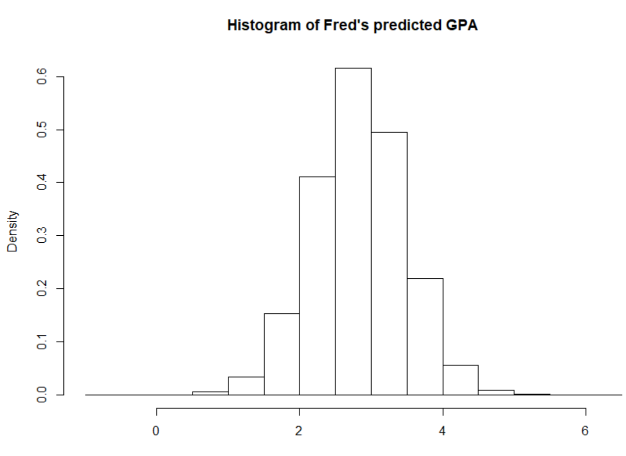
I have $Y \sim \mathrm{bin}(25, \theta)$. This sample $Y = 0$. Then

$$\begin{aligned} p(\theta|y) &\propto p(y=0|\theta)p(\theta) \\ &\propto \theta^y(1-\theta)^{25-y}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &\propto \theta^{\alpha-1}(1-\theta)^{25+\beta-1} \\ \{\theta|y=0\} &\sim \mathrm{beta}(\alpha, 25+\beta) \end{aligned}$$

Point estimate of θ is $\hat{\theta} = \mathrm{E}[\theta|y=0] = \alpha/(25+\beta) = 0.014\,845\,36$. 95% upper bound for θ is $\mathrm{qbeta}(.95, \alpha, 25+\beta) = 0.048\,569\,43$, giving interval for θ is $(0, 0.048\,569\,43)$.

2

Took $S = 1\,000\,000$ Monte Carlo samples of posterior distribution of β. Marginal posterior means $\hat{\beta}_1 = 0.016\,699\,77$, $\hat{\beta}_2 = 0.020\,701\,44$. Marginal 95% HPD intervals showed $\beta_1 \in (0.007\,630\,409, 0.025\,895\,26)$, $\beta_2 \in (0.011\,814\,688, 0.029\,731\,22)$.
95% HPD interval for $\mathrm{E}[Y^*|X_1^* = 70, X_2^* = 80] \in (2.613\,156, 3.033\,227)$.
95% HPD for $\{Y^*|X_1^* = 70, X_2^* = 80\} \in (1.576\,744, 4.090\,318)$. Used the Monte Carlo sample as the posterior distribution of $\{\beta_1, \beta_2, \sigma^2|Y\}$ and found 95% HPD using `HPDinterval`. Predicted Mary's GPA from each Monte Carlo



sample and subtracted Fred's predicted GPA from the same sample. The proportion of these that were greater than 0 are the probability of Mary having a higher GPA and is equal to $0.555\,257$.

3

Comparing the data in `menchild30nobach.dat` to a new mixture model. The sampling and prior distributions for the mixture model are

$$
\begin{aligned}
\{Y_i|X_i,\theta\} &\sim \begin{cases} 0 & \text{individual } i \text{ is unmarried}(X_i=0) \\ \text{Poisson}(\theta) & \text{individual } i \text{ is married}(X_i=1) \end{cases} \\
\{X_i|p\} &\sim \text{Bernoulli}(p) \\
p &\sim \text{Beta}(1,1) \\
\theta &\sim \text{Gamma}(1,\frac{1}{3}).
\end{aligned}
$$

The prior on p was selected as a minimal prior, while the prior on θ loosely represented my idea that families have 2-3 children on average. The goal is to see if this model fits the data better than a Poisson model. In order to sample from the new model, I made a Gibbs sampler. The full conditional distributions derived for p and θ are

$$
\begin{aligned}
p(p|y,x,\theta) &\propto p(x|p)p(p) \\
&\propto \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&\propto p^{\sum x_i}(1-p)^{n-\sum x_i} \\
\{p|Y,X,\theta\} &\sim \text{Beta}\left(\sum_{i=1}^{n} x_i+1, n-\sum_{i=1}^{n} x_i+1\right)
\end{aligned}
$$

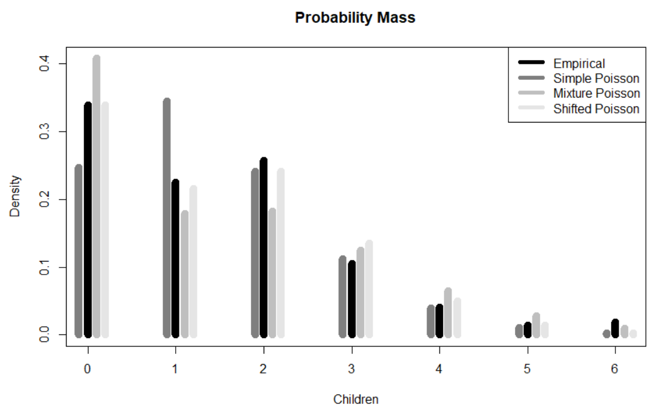
$$
\begin{aligned}
p(\theta|y,x,p) &\propto p(y|x,\theta)p(\theta) \\
&\propto \prod_{\{i|x_i=1\}} (\theta^{y_i}e^{-\theta})e^{-\frac{\theta}{2}} \\
&\propto \theta^{n_1\bar{y}_1}e^{-(n_1+\frac{1}{2})\theta} \\
\{\theta|y,x,p\} &\sim \text{Gamma}\left(n_1\bar{y}_1+1, n_1+\frac{1}{2}\right).
\end{aligned}
$$

For x, I used basic probability theory,

$$
\begin{aligned}
p(x_i=0,y_i=0|\theta,p) &= p & p(x_i=0,y_i>0|\theta,p) &= 0 \\
p(x_i=1,y_i=0|\theta,p) &= (1-p)e^{-\theta} & p(x_i=1,y_i>0|\theta,p) &= (1-p)(1-e^{-\theta})
\end{aligned}
$$

$$
\begin{aligned}
p(x_i=0|y_i=0,\theta,p) &= \frac{p}{p+(1-p)\exp(-\theta)} & p(x_i=0|y_i>0,\theta,p) &= 0 \\
p(x_i=1|y_i=0,\theta,p) &= \frac{(1-p)\exp(-\theta)}{p+(1-p)\exp(-\theta)} & p(x_i=1|y_i>0,\theta,p) &= 1.
\end{aligned}
$$

I get a Bayes factor of 56.132 82 (log 4.027 721) to show a preference for the new mixture model over the Poisson model, leading me to prefer the new model, although graphically it is not ideal. The mixture model mean is 1.403 08, close to the sample mean of 1.399 083, with a variance of 2.351 91, not as far from the sample variance 1.918 34 as the simple Poisson model (also not as far in standard deviation). I also decided to check a similar (and simpler) mixture model, where $Y=0$ with probability p, otherwise $Y-1 \sim \text{Poisson}(\theta)$. This shifted Poisson model gave much better results visually, and had a Bayes factor of 81.832 72 (log 4.404 677) against the mixture model.



4 Appendix

```
#### Math 457 ####
#### Final Exam ####
#### Miles Keppler ####

setwd("U:/Winter␣2018/M457/")
library(coda)

#### 1 ####

mean.theta = .03
sd.theta = .07/2
beta = mean.theta / sd.theta^2
alpha = beta * mean.theta

alpha/(25+beta)

qbeta(.95, alpha, beta+25)

curve(dbeta(x,alpha,beta),lty=2,col="gray60")
curve(dbeta(x,alpha,beta+25),add=T)
legend("topright",legend=c("Prior","Posterior"),lty=c(2,1),col=c("gray60","black"))

#### 2 ####

# gprior function from
# http://wwwlegacy.stat.washington.edu/people/pdhoff/Book/ComputerCode/regression_gprior.r
lm.gprior<-function(y,X,g=dim(X)[1],nu0=1,
                    s20=try(summary(lm(y~-1+X))$sigma^2,silent=TRUE),S=1000) {
  n<-dim(X)[1] ; p<-dim(X)[2]
  Hg<- (g/(g+1))*X%*%solve(t(X)%*%X)%*%t(X)
  SSRg<- t(y)%*%(diag(1,nrow=n)-Hg)%*%y

  s2<-1/rgamma(S,(nu0+n)/2,(nu0*s20+SSRg)/2)

  Vb<- g*solve(t(X)%*%X)/(g+1)
  Eb<- Vb%*%t(X)%*%y

  E<-matrix(rnorm(S*p,0,sqrt(s2)),S,p)
  beta<-t(t(E%*%chol(Vb))+c(Eb))
```

```
  list(beta=beta,s2=s2)
}

data = read.table("GPAdata.dat",header = T)
Y = data$y
X = cbind(data$x1,data$x2)
n = length(Y)
S = 1000000
myglm = lm.gprior(y = Y, X = X, g = n, nu0 = 2, s20 = .4^2, S = S)

colMeans(myglm$beta)
HPDinterval(as.mcmc(myglm$beta))

fredmean = 70*myglm$beta[,1]+80*myglm$beta[,2]
HPDinterval(as.mcmc(fredmean))

fredsample = fredmean+rnorm(S,0,sqrt(myglm$s2))
HPDinterval(as.mcmc(fredsample))
hist(fredsample,probability=T,main="Histogram of Fred's predicted GPA",xlab="GPA")

marysample = 65*myglm$beta[,1]+90*myglm$beta[,2]+rnorm(S,0,sqrt(myglm$s2))
sum(marysample>fredsample)/S

#### 3 ####

set.seed(111)
Y = scan("menchild30nobach.dat")

findMass = function(x, t) {
  n = as.numeric(names(t))
  for (i in 1:length(n)) {
    if (n[i] == x) {
      return(as.numeric(t)[i])
    }
  }
  return(0)
}

# Gibbs sampler (mixture model)
# should look like scaled poisson plus mass at 0

n = length(Y)
p = .5
theta = 2

M = 100000

PHI = matrix(NA, ncol = 2, nrow = M)
for (i in 1:M) {
  X = NULL
  for (j in 1:n) {
    if (Y[j]==0) {
      X[j] = rbinom(1,1,(1-p)*exp(-theta)/(p+(1-p)*exp(-theta)))
    } else {
      X[j] = 1
    }
```

```
  }
  sx = sum(X)
  p = rbeta(1,sx+1,n-sx+1)
  theta = rgamma(1,sum(Y[X==1])+1,sx + 1/3)
  PHI[i,] = c(theta, p)
}

NY = rep(NA, M)
for (i in 1:M) {
  theta = PHI[i,1]
  p = PHI[i,2]
  x = rbinom(1,1,p)
  ny = 0
  if (x==1) {
    ny = rpois(1, theta)
  }
  NY[i] = ny
}

# Plotting

# Empirical
taby = table(Y)/n
tabny = table(NY)/M
yind = as.numeric(names(taby))
plot(as.numeric(yind),c(taby),type="h",ylim = c(0,max(c(taby,tabny))),lwd=10,
     main = "Probability Mass", xlab = "Children", ylab = "Density", xlim = c(0-.1, max(yind)
# Mixture Poisson
nyind = as.numeric(names(tabny))
points(as.numeric(nyind)+.1,c(tabny),type="h",lwd=10,col="gray75")
# Poisson model
points(0:max(yind)-.1,dpois(min(yind):max(yind),mean(Y)),type="h",lwd=10,col="gray50")

p = sum(Y==0)/n
mytab = as.table(c(p,(1-p)*dpois(as.numeric(names(taby))[-1]-1,mean(Y[Y!=0]-1))))
names(mytab) = mytab.names = 0:(length(mytab)-1)
points(mytab.names+.2,mytab,type="h",lwd=10,col="gray90")

legend("topright", legend = c("Empirical","Simple Poisson","Mixture Poisson","Shifted Poisson
       lwd = c(5,5,5,5), col = c("black","gray50","gray75","gray90"))

LL = c(0,0,0)
for (i in 1:n) {
  LL[1] = LL[1] + log(findMass(Y[i], tabny))
  LL[2] = LL[2] + log(dpois(Y[i],mean(Y)))
  LL[3] = LL[3] + log(findMass(Y[i], mytab))
}
bf = c(exp(LL[1]-LL[2]), exp(LL[3]-LL[1]))
bf ; bf/(1+bf)

nLL = 0
for (i in 1:n) {
  nLL = nLL + log(findMass(Y[i],taby))
}
```