ISYE 6420 Bayesian Statistics

Homework 6

November 22th, 2021

1.

Provided is the code section for the regression formulation building off of the gastro example which uses the weib function to calculate our updated times[i] at each iteration followed by the calculations for lambda and S, in order to calculate the tumor profile as covariate.

We will burn the first 1000 samples and generate statistic for the next 100,000 samples with the following code:

```
 model
 {
     for(i in 1 : N) {
      times[i] ~ dweib(v,lambda[i])I(censor[i],)
      lambda[i] <- exp(beta0 + beta1*type[i])
      s_value[i]  <- exp(-lambda[i]*pow(times[i],v));
          f[i]  <- lambda[i]*v*pow(times[i],v-1)*s_value[i]
          h[i] <- f[i]/s_value[i]
          index[i] <- i
    }
    v ~ dexp(0.001)
    beta0 ~ dnorm(0.0, 0.0001)
    beta1 ~ dnorm(0.0, 0.0001)
   }

DATA
list(N=80)

 TonguesData

INITS
list(v = 1, beta0 = 0, beta1=0)
```

The results in the following output:

Node statistics

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta0	-3.195	0.5634	0.02517	-4.329	-3.177	-2.132	1001	100000
beta1	-0.5557	0.2852	0.01099	-1.105	-0.5523	-0.01844	1001	100000
v	0.8449	0.09834	0.003365	0.6608	0.8409	1.048	1001	100000

Figure 1: Summary Statistics

We can see that for the beta1 slope this results in the following 95% credible set of: $[-1.105, -0.5523]$

2.

Provided is the code section used to run poisson regression to predict against Y:

```
# MODEL
model{
for( i in 1:N ) {
y[i] ~ dpois(mu[i])
mu[i] <- exp( beta0 + beta1*  x[i])
}
# setting the priors for beta0 and 1
beta0 ~ dnorm(0, 0.001)
beta1 ~ dnorm(0, 0.001)

# iterate through observations and predict missing observations
using
# possion regression hint provided
for(i in 1:N) {
x[i] ~ dpois(2)
}

#Get the average # broken packages X = 4
mean_ypred <- exp( beta0 + beta1*4)
 #Use poisson regression to get the prediction result of # broken
packages at X = 4
ypred ~ dpois(mean_ypred)
}
```

```
# DATA
list(N=15,
x = c(2, 1, 0, 2, NA, 3, 1, 0, 1, 2, 3, 0, 1, NA, NA),
y =c(NA, 16, 9, 17, 12, 22, 13, 8, NA, 19, 17, 11, 10, 20, 2))

# INIT
list(beta0=2,beta1=0)
```

Then this is the summary statistics after burning the first 1000 observations.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta0	2.17	0.1433	0.00285	1.881	2.174	2.44	1001	100000
beta1	0.2891	0.07437	0.001485	0.1421	0.2878	0.4367	1001	100000
deviance	109.6	3.288	0.03326	105.5	108.7	117.8	1001	100000
mean_ypred	28.38	5.639	0.09077	18.72	27.84	40.97	1001	100000
x[5]	1.276	0.7843	0.003162	0.0	1.0	3.0	1001	100000
x[14]	2.549	0.827	0.004045	1.0	3.0	4.0	1001	100000
x[15]	0.1958	0.4302	0.002627	0.0	0.0	1.0	1001	100000
y[1]	15.68	4.184	0.01634	8.0	15.0	24.0	1001	100000
y[9]	11.73	3.591	0.02014	5.0	12.0	19.0	1001	100000
ypred	28.37	7.726	0.09013	15.0	28.0	45.0	1001	100000

Figure 2: Summary Statistics

Now we will go through the each part from a-d:

a)
When we set deviance in the sampling tab we get the following results in openbugs

$mean\ deviance\ = 109.6$
$This\ produces\ a\ 95\%\ credible\ set\ of\ [105.5,\ 117.8]$

b)
The number packages on average that are expected will be broken if the number of shipment routes is X = 4 is measured by the $mean\ ypred$ variable, with a 95% credible set of $[18.72,\ 40.97]$

c)
For a particular shipment sent from Shenzhen that involves X = 4 shipping routes. $ypred$ represents the predicted number of broken packages. This variable has the following outputs:
$95\%\ credit\ set\ = [15,\ 45]$
$bayesian\ estimate\ = \ 28.38$

We can see the differences from b that the credible set in c contains all of the values and covers the full range of the credible set of b.

d)

The estimates of the unobserved provided by the mean value are:

$X_5 = 1.276$

$X_{14} = 2.549$

$X_{15} = 0.1958$

$Y_1 = 15.68$

$Y_9 = 11.73$