

Computer Lab 4

Andrea Bruzzone, Thomas Zhang

2016 M05 18

Assignment 1

a,

We do a `glm()` fit and obtain the maximum likelihood estimator of β in the Poisson regression model for the eBay data.

```
## [1] "The maximum likelihood estimator of beta coefficients:"
```

```
## PowerSeller    VerifyID      Sealed      Minblem      MajBlem      LargNeg
## 0.71896702 -0.13883985  0.71681438  0.37324148  0.01400166  0.64177063
##      LogBook MinBidShare
## -0.22946925 -3.23450853
```

Through the `summary.glm()` function, we can say that it looks like VerifyID, Sealed, MajBlem, LogBook and MinBidShare are significant predictors in this model.

Next, we do a Bayesian analysis of the Poisson regression with prior distribution $\beta \sim \mathcal{N}(\mathbf{0}, 100 \cdot (X'X)^{-1})$. We know that we can use the `optim()` function to numerically find the posterior mode $\tilde{\beta}$ and the Hessian $-J_{y,\tilde{\beta}}$ at that posterior mode. With these values we can approximate the posterior distribution as a multivariate normal distribution, $\beta|y \sim \mathcal{N}(\tilde{\beta}, J_{y,\tilde{\beta}}^{-1})$.

```
## [1] "The posterior mode beta coefficients: "
```

```
## [1] 0.72028328 -0.13469969 0.71574282 0.37329126 0.01443572 0.64240277
## [7] -0.22694135 -3.21879090
```

```
## [1] "The hessian at the posterior mode: "
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] -1576.80764 -68.65215 -3.818716e+02 -1.378046e+02 -9.355086e+01
## [2,]  -68.65215 -128.73447 -5.880810e+01 -1.554736e+01  0.000000e+00
## [3,] -381.87165 -58.80810 -5.057035e+02 -5.684342e-08 -5.684342e-08
## [4,] -137.80459 -15.54736 -5.684342e-08 -3.091830e+02  0.000000e+00
## [5,]  -93.55086  0.00000 -5.684342e-08  0.000000e+00 -1.270093e+02
## [6,]  -57.11834  0.00000  5.684342e-08 -4.660851e+01 -5.684342e-08
## [7,] -103.04136 -81.57867 -1.618865e+02 -6.939475e+01 -3.779654e+01
## [8,]  304.58623  40.44637  1.193921e+02  7.699874e+01  3.972684e+01
##           [,6]      [,7]      [,8]
## [1,] -5.711834e+01 -103.04136  304.58623
## [2,]  0.000000e+00 -81.57867  40.44637
## [3,]  5.684342e-08 -161.88649  119.39211
## [4,] -4.660851e+01 -69.39475  76.99874
## [5,] -5.684342e-08 -37.79654  39.72684
## [6,] -3.858052e+02 -215.89732  122.08186
## [7,] -2.158973e+02 -1565.97023  444.17956
## [8,]  1.220819e+02  444.17956 -354.63256
```

```
## [1] "A posterior draw of beta: "
```

```
##          [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
## [1,] 0.7111967 -0.2890974 0.8119556 0.35097 -0.01587766 0.6786866
##          [,7]      [,8]
## [1,] -0.2393631 -3.331718
```

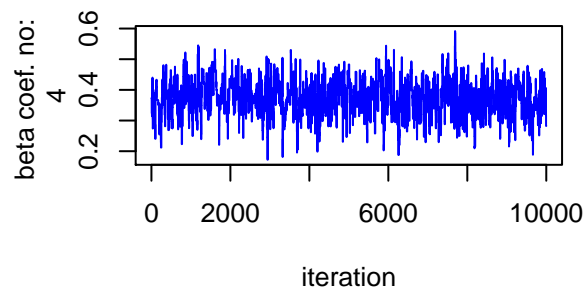
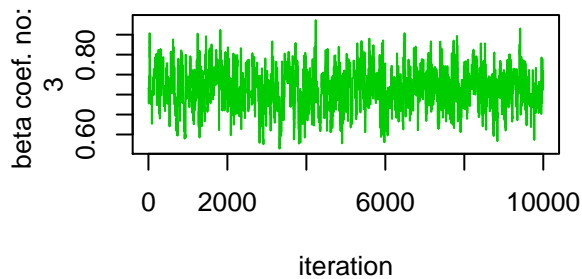
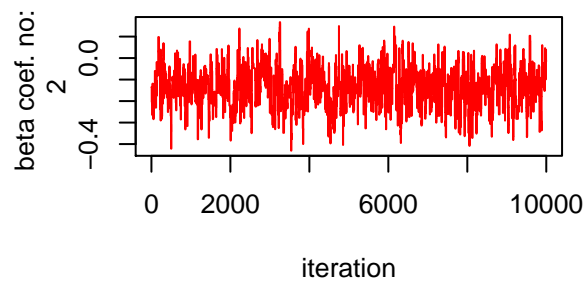
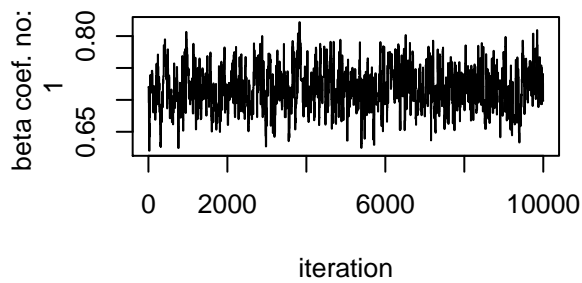
Now we simulate from the actual posterior of β using the random walk Metropolis-Hastings algorithm. We are going to use a multivariate normal density, $\theta_p|\theta_c \sim \mathcal{N}(\theta_c, \tilde{c} \cdot \Sigma)$ as proposal density where $\Sigma = J_{y,\tilde{\beta}}^{-1}$ and \tilde{c} is equal to 2.4 divided by the square root of the number of parameters.

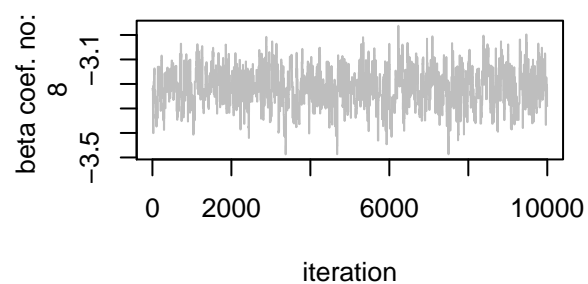
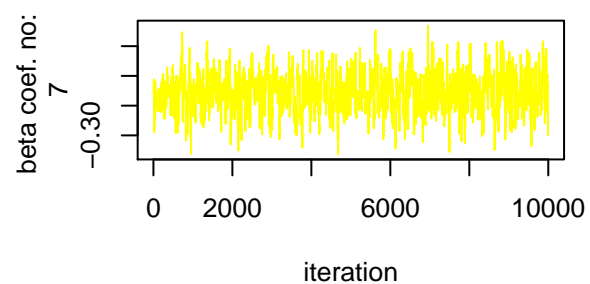
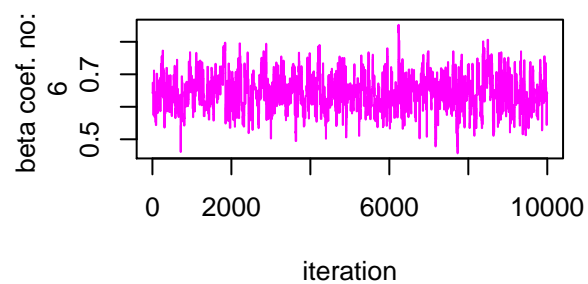
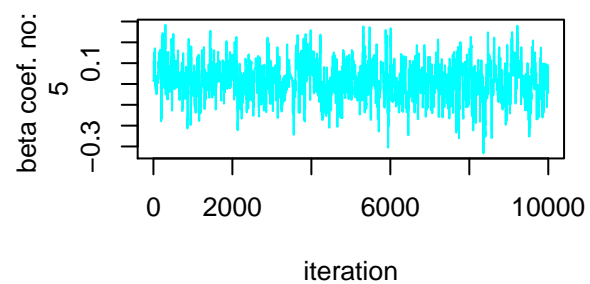
```
## [1] "last iteration of M-H algorithm: "
```

```
## [1] 0.69845754 -0.01422904 0.75913661 0.28230703 0.05263073 0.64118339
## [7] -0.29970627 -3.28996248
```

As it can be seen the draw from the M-H algorithm is very similar to the one found before.

We plot the traceplots for each beta coefficient and then we plot the histograms for the posterior distributions of $\phi_j = \exp \beta_j$.

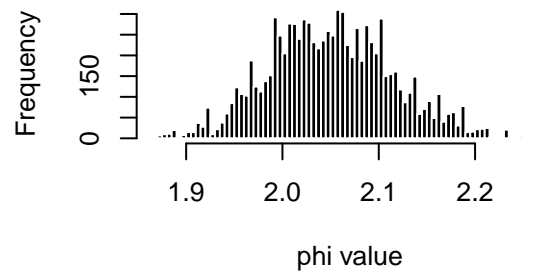




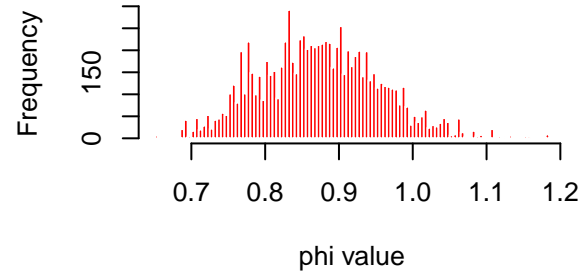
```
## [1] "Effective sizes of the MCMC chains, 10000 iterations: "
```

```
##      var1      var2      var3      var4      var5      var6      var7      var8
## 344.2819 370.3795 419.2641 413.6408 397.9017 355.0700 396.7633 374.3520
```

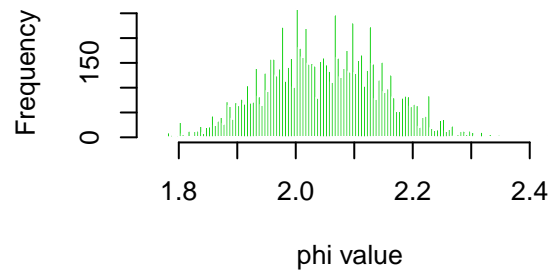
Posterior dist. of phi coef. no:
1



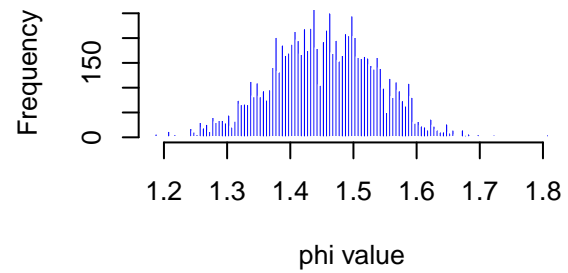
Posterior dist. of phi coef. no:
2

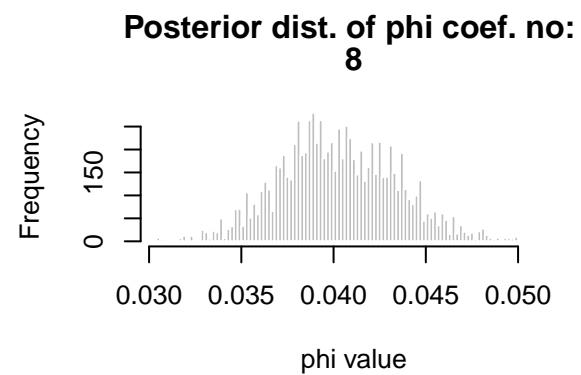
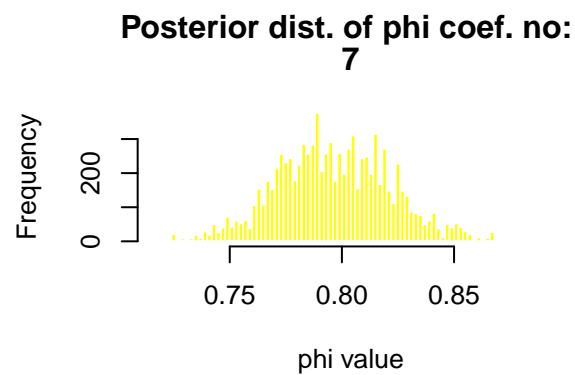
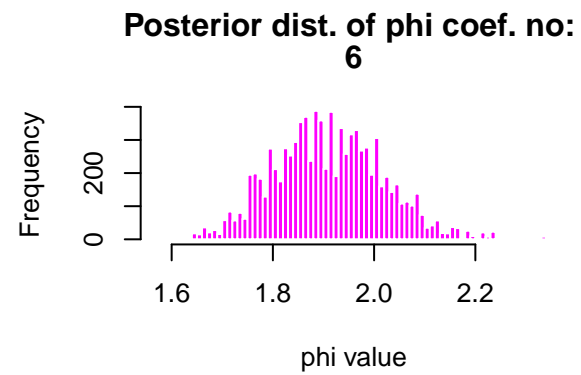
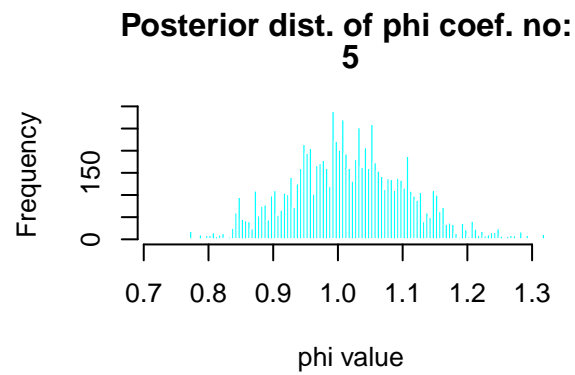


Posterior dist. of phi coef. no:
3



Posterior dist. of phi coef. no:
4

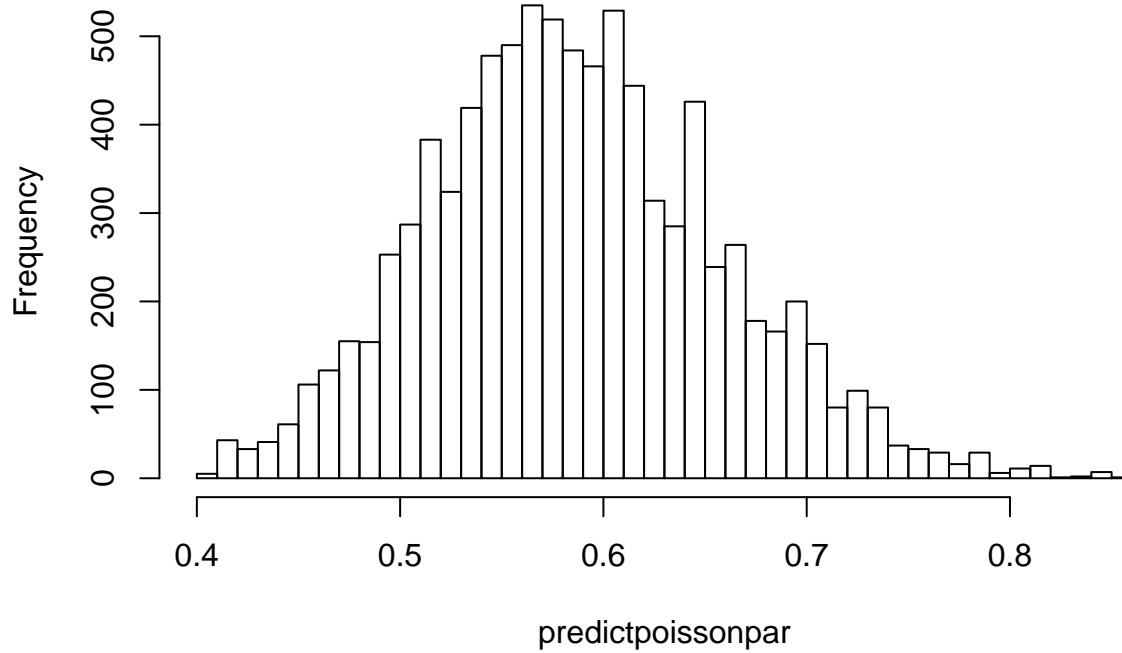




From the traceplots we can see that all the β values seem to converge. Anyways the effective samples size is only about 3.5%.

Finally we plot a histogram over the predictive distribution of poisson parameter λ for the auction given in the lab instructions and calculate the probability that that auction will have zero bids.

Histogram of predictpoissonpar



```
## [1] "Mean probability for no bids on the new auction: 0.558"
```

We can tell that the coin object and the seller is of good quality, but that the MinBidShare and LogBook values were rather higher than average. These factors balance each other out and we believe that the result is reasonable.