

STAT 5370 – Decision Theory

Homework 1

Matthew Aaron Looney

6/9/2018

Contents

| | |
|--|-----------|
| Problem 1 | 2 |
| (a) For the Poisson model, identify the conjugate prior. This should be a general class of priors. | 2 |
| (b) Under the conjugate prior, derive the posterior distribution of λy . This should be a general expression based on the choice of the hyper-parameters specified in your prior. | 2 |
| (c) Find the posterior mean and variance of λy . These should be general expressions based on the choice of the hyper-parameters specified in your prior. | 3 |
| (d) Obtain the MLE of λ . Develop and discuss a relationship that exists between the MLE and posterior mean identified in (c). . | 4 |
| (e) Write two separate R programs which can be used to find both a $(1-\alpha)100\%$ equal-tailed credible interval and a $(1-\alpha)100\%$ HPD credible interval for the Poisson model. These programs should take as arguments the following inputs: the observed data, prior hyper-parameters, and significance level. | 5 |
| (f) Find a data set which could be appropriately analyzed using the Poisson model. This data set should be of interest to you, and you should discuss, briefly, why the aforementioned model is appropriate; e.g., consider independence, identically distributed, etc. etc. You will also need to provide the source of the data. . | 5 |
| (g) Analyze the data set you have selected in (e). Provide posterior point estimates of λ , credible intervals, etc. etc. Your analysis should be accompanied by an appropriate discussion of your findings. | 7 |
| Appendix A - Code for Problem 1, part e. | 9 |
| Appendix B - Data | 13 |

Problem 1

In this problem we will consider developing a Bayesian model for Poisson iid data; i.e., our observed data will consist of $Y_1, \dots, Y_n \stackrel{iid}{\sim} \text{Poisson}(\lambda)$. Recall, a random variable Y is said to follow a Poisson distribution, with mean parameter λ , if its pmf is given by

$$p(y|\lambda) = \frac{e^{-\lambda} \lambda^y}{y!} I(y \in \{0, 1, 2, \dots\}) \quad (1)$$

Note, the Poisson model is often used to analyze count data.

(a) For the Poisson model, identify the conjugate prior. This should be a general class of priors.

If we assume the Gamma distribution as a prior for the Poisson model: $p(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \propto \lambda^{\alpha-1} e^{-\beta\lambda}$; where $\alpha > 0$ (shape parameter) and $\beta > 0$ (rate parameter) then this choice of prior is also the conjugate prior for the Poisson model.

(b) Under the conjugate prior, derive the posterior distribution of $\lambda|y$. This should be a general expression based on the choice of the hyper-parameters specified in your prior.

To derive the posterior distribution we need the likelihood function associated with the Poisson model:

$$L(\lambda|y) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!} \quad (2)$$

In general, the posterior distribution is defined as:

$$p(\lambda|y) = \frac{p(y|\lambda)p(\lambda)}{\int_A p(y|\lambda)p(\lambda)d\lambda} \propto \underbrace{p(y|\lambda)}_{L(\lambda)} \underbrace{p(\lambda)}_{\text{Prior}} \quad (3)$$

With a Poisson DGP and a Gamma prior the derived posterior distribution is defined by:

$$\begin{aligned} p(\lambda|y) &= e^{-n\lambda} \lambda^{\sum_{i=1}^n y_i} \cdot \lambda^{\alpha-1} e^{-\beta\lambda} \\ &= \lambda^{\sum y_i + \alpha - 1} \cdot e^{-(n+\beta)\lambda} \end{aligned} \quad (4)$$

The posterior distribution of λ under the Gamma prior is $Gamma(\sum y_i + \alpha, n + \beta)$, where α and β are the hyperparameters of the Gamma distribution.

(c) Find the posterior mean and variance of $\lambda|y$. These should be general expressions based on the choice of the hyper-parameters specified in your prior.

The mean and variance of the Gamma distribution is given by the following expressions:

$$E(\lambda) = \frac{\alpha}{\beta}; \text{Var}(\lambda) = \frac{\alpha}{\beta^2}.$$

Then the posterior mean and variance of $\lambda|y$ is:

$$\begin{aligned} E(\lambda|y) &= \frac{\alpha}{\beta} = \frac{\sum y_i + \alpha}{n + \beta} \\ \text{Var}(\lambda|y) &= \frac{\alpha}{\beta^2} = \frac{\sum y_i + \alpha}{(n + \beta)^2} \end{aligned} \quad (5)$$

(d) Obtain the MLE of λ . Develop and discuss a relationship that exists between the MLE and posterior mean identified in (c).

Starting from the Poisson distribution, $p(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$, we obtain the likelihood function:

$$L(\lambda|y_1, y_2, \dots, y_n) = \prod_{i=1}^n \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} \quad (6)$$

Then, taking log of equation 6 and differentiating w.r.t λ we can obtain an expression for $\hat{\lambda}$:

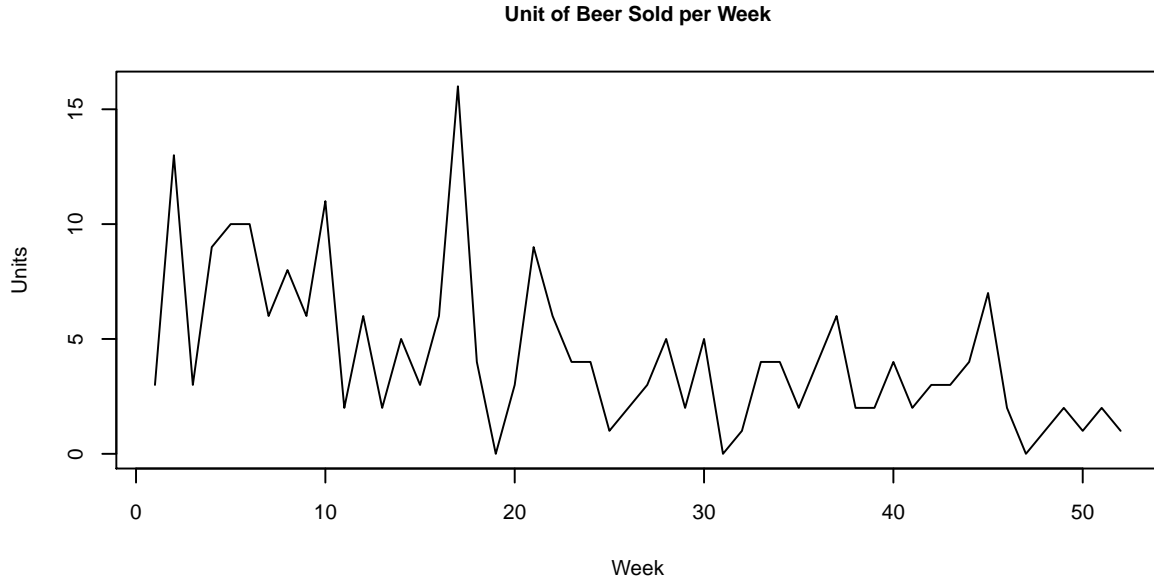
$$\begin{aligned} \log L(\lambda|x_1, x_2, \dots, x_n) &= -n\lambda - \sum_{i=1}^n \log(y_i!) + \log(\lambda) \sum_{i=1}^n y_i \\ \frac{\partial \log L(\cdot)}{\partial \lambda} &= -n + \frac{\sum_{i=1}^n y_i}{\hat{\lambda}} \stackrel{set}{=} 0 \\ \hat{\lambda} &= \frac{\sum_{i=1}^n y_i}{n} \end{aligned} \quad (7)$$

We can now see that when the hyperparameters of the posterior mean (part c, equation 5) are zero the posterior mean collapses to the MLE estimator of λ . This shows that the MLE estimator of λ and the posterior mean of λ are related by a weighted average. Further, we can see that the MLE estimator for λ is $\sum_{i=1}^n y_i/n$ and the posterior mean are a combination of the weighted average of $\sum_{i=1}^n y_i/n$ and the prior mean.

(e) Write two separate R programs which can be used to find both a $(1-\alpha)100\%$ equal-tailed credible interval and a $(1-\alpha)100\%$ HPD credible interval for the Poisson model. These programs should take as arguments the following inputs: the observed data, prior hyper-parameters, and significance level.

Please see Appendix A

(f) Find a data set which could be appropriately analyzed using the Poisson model. This data set should be of interest to you, and you should discuss, briefly, why the aforementioned model is appropriate; e.g., consider independence, identically distributed, etc. etc. You will also need to provide the source of the data.



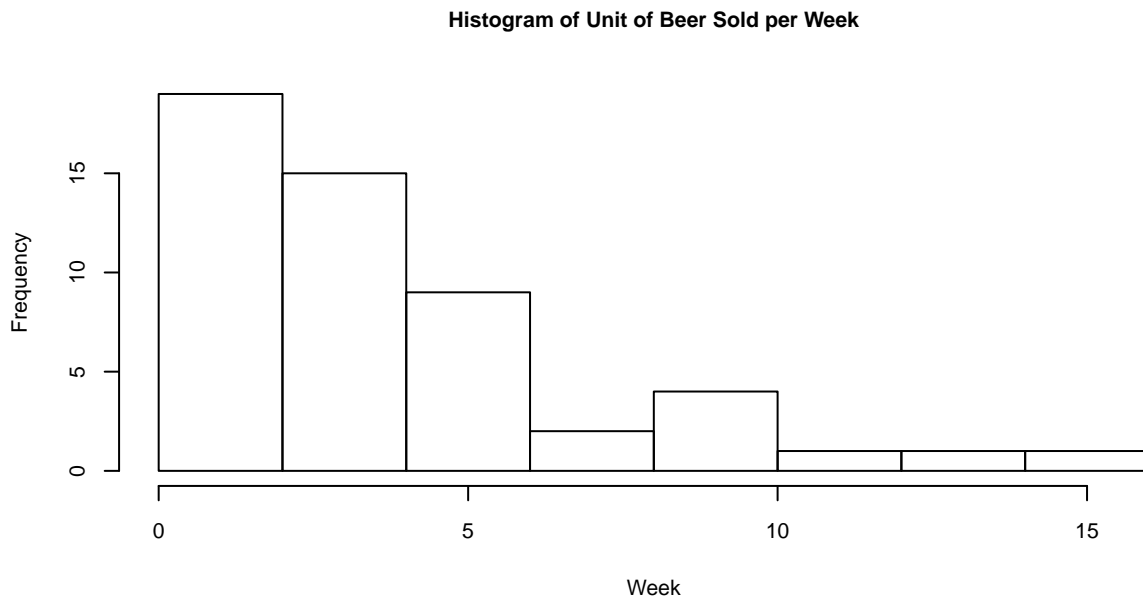
In this study I will use a Poisson model to analyze market level data for beer. A market is defined in the follow way: units of beer sold per week during 2010 in Los Angeles, California over a 52 week period at the chain/brand level. The market level

sales data are obtained from Information Resource Inc. (IRI) and consist of all dollar by volume sales for a brand of beer at a specific chain during a one week period. In this data set there are seven chains represented in Los Angeles; I have chosen a single chain for this analysis. The chain name is masked for privacy. The brand is a California craft brew named Anchor Steam Beer. The brewery is located in San Francisco, California and has a total annual production output of 132,000 barrels.

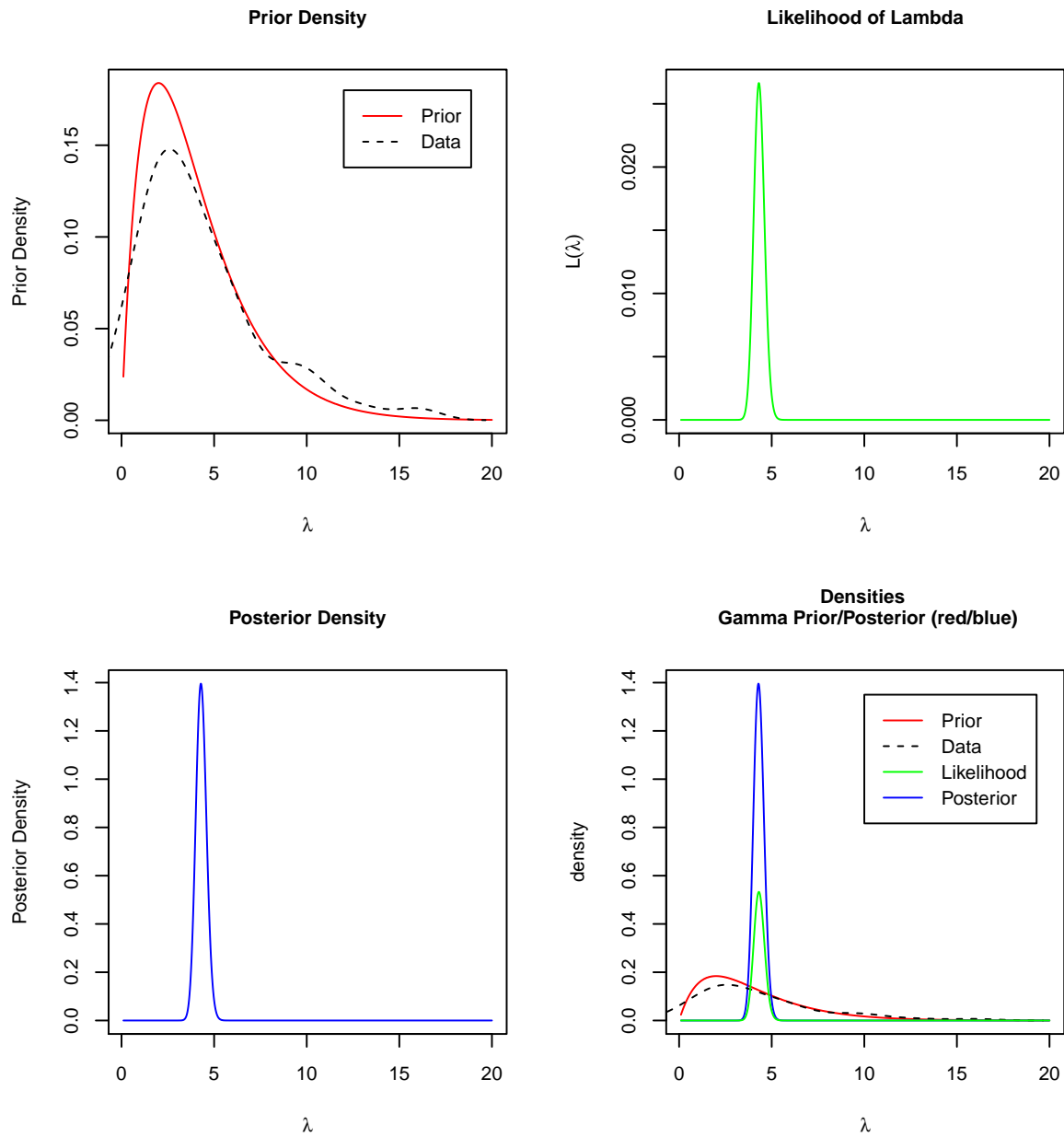
Given the count nature of this data, using a Poisson model is appropriate. Assuming independent and identically distributed data is reasonable since there is no expectation that aggregate beer sales from one week to the next should be related in any meaningful way. Of course this assumption is not very hard to counter with some thought but I will assume iid data in this analysis.

A single unit of beer consists of 6 - 12 ounce bottles of Anchor Steam Beer.

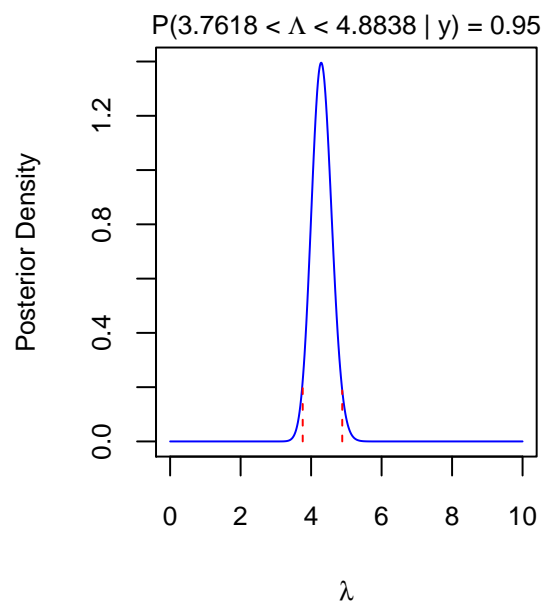
See Appendix B for data.



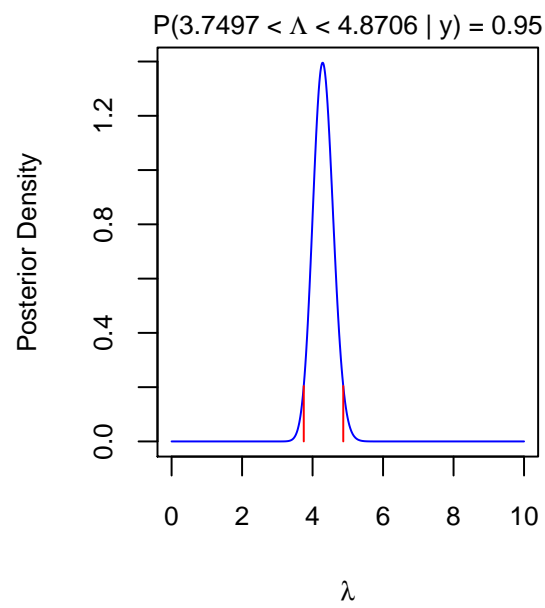
(g) Analyze the data set you have selected in (e). Provide posterior point estimates of λ , credible intervals, etc. etc. Your analysis should be accompanied by an appropriate discussion of your findings.



Plot of Equal Tail Credible Interval



Plot of HPD Credible Interval



Appendix A - Code for Problem 1, part e.

```
# Equal Tail Credible Interval -----  
equalTail_CI <- function(y, a0, b0, alpha=0.05, displayPlot=1){  
  
  n= length(y)  
  a_posterior= a0+ n* mean(y)  
  b_posterior= b0+ n  
  critI <- qgamma(p=c(0.025, 0.975), shape= a0+ n* mean(y), rate= b0+ n)  
  
  if(displayPlot==1){  
    th = seq(0, 10, length=1000)  
  
    plot(th, dgamma(th, a_posterior, b_posterior),  
         t="l", lty=1,xlab=expression(lambda),  
         ylab="Posterior Density",  
         col="blue")  
    #abline(h= h*dmode)  
    segments(critI[2], 0, critI[2], dgamma(critI[2], a_posterior, b_posterior),  
            col="red", lty=2)  
    segments(critI[1], 0, critI[1], dgamma(critI[1], a_posterior, b_posterior),  
            col="red", lty=2)  
  
    title(main="Plot of Equal Tail Credible Interval")  
  }  
}
```

```

      mtext(bquote(paste("P(", .(round(critI[1], 4)), " < ", Lambda, " < ",
                          .(round(critI[2], 4)), " | " , y, ") = ",
                          .(round(0.95, 4)))
            ), 3, line=0)

    } else {}

  return(critI)

}

# HDP Credible Interval -----

HPD.h = function(y, h=.1, a_posterior=a1, b_posterior=b1, displayPlot=0, ...){

  if(a_posterior >= 1){
    mode = (a_posterior - 1) / (b_posterior)
    dmode = dgamma(mode, a_posterior, b_posterior)
  } else {
    print("Can not calculate mode.
          The b_posterior must be greater than or equal to zero!")
  }

  lt = uniroot(f=function(x){
    dgamma(x, a_posterior, b_posterior) / dmode - h},
    lower=0,
    upper=mode)$root

```

```

ut = uniroot(f=function(x){
  dgamma(x,a_posterior, b_posterior) / dmode - h},
  lower=mode,
  upper=100)$root

coverage = pgamma(ut, a_posterior, b_posterior) -
  pgamma(lt, a_posterior, b_posterior)

if(displayPlot==1){
  th = seq(0, 10, length=1000)

  plot(th, dgamma(th, a_posterior, b_posterior),
    t="l", lty=1,xlab=expression(lambda),
    ylab="Posterior Density",
    col="blue")

  #abline(h= h*dmode)

  segments(ut, 0, ut, dgamma(ut, a_posterior, b_posterior), col="red")
  segments(lt, 0, lt, dgamma(lt, a_posterior, b_posterior), col="red")

  title(main="Plot of HPD Credible Interval")

  mtext(bquote(paste("P(", .(round(lt, 4)), " < ", Lambda, " < ",
    .(round(ut, 4)), " | " , y, ") = ",
    .(round(coverage, 4)))
    ), 3, line=0)

} else {}

```

```
    return(c(lt, ut, coverage, h))

}

HPD = function(h, y, alpha){
  cov = HPD.h(y, h = h)[3]
  res = (cov - (1 - alpha))^2
  return(res)
}
```

Appendix B - Data

Table 1: Data

| week | units | Calendar.week.starting.on |
|------|-------|---------------------------|
| 1 | 3 | 2009-12-28 |
| 2 | 13 | 2010-01-04 |
| 3 | 3 | 2010-01-11 |
| 4 | 9 | 2010-01-18 |
| 5 | 10 | 2010-01-25 |
| 6 | 10 | 2010-02-01 |
| 7 | 6 | 2010-02-08 |
| 8 | 8 | 2010-02-15 |
| 9 | 6 | 2010-02-22 |
| 10 | 11 | 2010-03-01 |
| 11 | 2 | 2010-03-08 |
| 12 | 6 | 2010-03-15 |
| 13 | 2 | 2010-03-22 |
| 14 | 5 | 2010-03-29 |
| 15 | 3 | 2010-04-05 |
| 16 | 6 | 2010-04-12 |
| 17 | 16 | 2010-04-19 |
| 18 | 4 | 2010-04-26 |
| 19 | 0 | 2010-05-03 |
| 20 | 3 | 2010-05-10 |
| 21 | 9 | 2010-05-17 |
| 22 | 6 | 2010-05-24 |
| 23 | 4 | 2010-05-31 |
| 24 | 4 | 2010-06-07 |
| 25 | 1 | 2010-06-14 |
| 26 | 2 | 2010-06-21 |
| 27 | 3 | 2010-06-28 |
| 28 | 5 | 2010-07-05 |
| 29 | 2 | 2010-07-12 |
| 30 | 5 | 2010-07-19 |
| 31 | 0 | 2010-07-26 |
| 32 | 1 | 2010-08-02 |
| 33 | 4 | 2010-08-09 |
| 34 | 4 | 2010-08-16 |
| 35 | 2 | 2010-08-23 |
| 36 | 4 | 2010-08-30 |
| 37 | 6 | 2010-09-06 |
| 38 | 2 | 2010-09-13 |
| 39 | 2 | 2010-09-20 |
| 40 | 4 | 2010-09-27 |
| 41 | 2 | 2010-10-04 |
| 42 | 3 | 2010-10-11 |
| 43 | 3 | 2010-10-18 |
| 44 | 4 | 2010-10-25 |
| 45 | 7 | 2010-11-01 |
| 46 | 2 | 2010-11-08 |
| 47 | 0 | 2010-11-15 |
| 48 | 1 | 2010-11-22 |
| 49 | 2 | 2010-11-29 |
| 50 | 1 | 2010-12-06 |
| 51 | 2 | 2010-12-13 |
| 52 | 1 | 2010-12-20 |

Units of Anchor Steam Beer sold in Los Angeles, CA (2010)