

# BST 140.778 Assignment 2

January 31, 2005

**Fine print** All computing assignments must be completed in the R statistical programming language. For non-computing assignments please appropriately typeset using  $\LaTeX$ . Bundle your R functions and  $\LaTeX$  files for each assignment in a zip or tar.gz file and email them to bcarvalh@jhsph.edu. Include *specific* instructions on how to run the code to answer the assignments in a README file. All code should be readable, formatted well, commented and clearly indicate the author and date. The general rule is: the more thought that has to go into understanding how to implement your code, the worse your grade will be. Please feel free to give each other small hints, but otherwise students must complete assignments individually.

1. Write a function `hpdInterval` that takes in count data and returns a HPD interval of size  $1 - \alpha$  calculated via bisection. Assume the Bayesian model is  $y_i|\mu \sim \text{Poisson}(\mu)$  and  $\mu \sim \text{Gamma}(\nu, \tau)$  with mean  $\nu/\tau$ . Test your program out on simulated data.

Function: `hpdInterval`

Inputs : a vector of counts, `y`, and size, `alpha`

Outputs : HPD interval for Bayesian model computed  
via bisection

2. Consider the counts in Table 1. Let  $y_i$  be the number of homicides that person  $i$  witnessed, for  $i = 1, \dots, 159$ . Assume each  $y_i|\theta_i$  is distributed as Poisson with mean  $\theta_i$ . Further assume the  $\theta_i$  are independently distributed as  $\text{Gamma}(\alpha, \beta)$ . Derive the marginal distribution of the  $y_i$ ; what distribution is it? Write a program that uses Newton Raphson to estimate  $(\alpha, \beta)$  using the marginal distribution of the  $y_i$ .

Function : `analyzeHomicide`

Inputs : none (the goal is just to analyze this data)

Outputs : List containing the MLE for alpha,beta and the  
2x2 observed information matrix

$y_i$	0	1	2	3	4	5
count	119	16	12	7	3	2

$y_i$  = response, number of homicides  
count = the number of times response  $y_i$  was observed

Table 1: Number of homicides witnessed in the past year.

3. Aitken's acceleration (see Burden and Faires, *Numerical Analysis*) is a way of speeding up a linearly convergent sequence. Suppose  $x_n \rightarrow x_\infty$  then Aitken noticed that

$$\frac{x_{n+1} - x_\infty}{x_n - x_\infty} \approx \frac{x_{n+2} - x_\infty}{x_{n+1} - x_\infty}.$$

This implies

$$x_\infty \approx x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}.$$

Aitken proved that the sequence defined by

$$\hat{x}_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}$$

converges more quickly than the  $x_n$ . Therefore Aitken's method is trivial to implement. One simply runs whatever algorithm generates the  $x_n$  normally and calculates the  $\hat{x}_n$  in progress; stop the algorithm when the change in the  $\hat{x}_n$ s gets small and take the final  $\hat{x}_n$  as your estimate of  $x_\infty$ .

Prove Aitken's acceleration results in faster convergence than the initial sequence. That is, assume

$$\lim_{n \rightarrow \infty} \frac{x_n - x}{x_{n-1} - x} = \lambda$$

where  $\lambda > 0$ . (Hence the  $\{x_n\}$  converge linearly). Show that

$$\lim_{n \rightarrow \infty} \frac{\hat{x}_n - x}{x_n - x} = 0.$$