2. **Capital Bikers – almost 400 hubs, over 3500 bikes – Estimate θ, the true average number of bikes rented at a hub across the system.**

   (a) The Poison model, parameterized by θ, is a reasonable sampling model because we are interested in estimating a mean from count data, the mean count of bikes, which is reasonably finite, within a closed space, the Washington D.C. area.

   Here we have data on 392 hubs, if the hubs can be thought of as independent, then $\Sigma y_i \sim$ POISON($\Sigma\theta_i$), where y_i represents the count of the number of bikes rented at each of the ith hubs and θ_i is the true number of bikes rented at each of the ith hubs.

   For this problem, I will say that $\theta_i = \theta$.

   (b) The parameter to estimate with a Poison model is θ, which here represents the true mean number of bikes as well as the true variation in the number of bikes. Counts of bikes will only be positive, as well as the mean and the variance is always positive. The Gamma model has a support of (0,) making it an ideal choice for the prior distribution on θ. It is also ideal because it weakens the direct mean-variance relationship seen in the Poison model by introducing an extra parameter in to the pdf.

   (c) If $\underset{\sim}{y}|\theta \sim$ POISON(θ), $\theta \sim$ GAMMA(α,β), then $\theta|\underset{\sim}{y} \sim$ GAMMA($\alpha + \Sigma y_i$, $\beta + n$).

   $\int \frac{\theta^{\Sigma y_i} e^{-n\theta}}{\Sigma y_i!} \times \frac{\beta^\alpha e^{-\beta\theta}\theta^{\alpha-1}}{\Gamma(\alpha)} d\theta$

   $\propto$

R Code

7. (a)

   (c)

   (d)