Chapter 13.3 Negative Binomial Sampling

Jim Albert and Monika Hu

Chapter 13 Case Studies

Introduction

- We presented evidence that the observed counts of "can" from a group of Federalist Papers of Alexander Hamilton were overdispersed.
- There was more variability in the counts than predicted by the Poisson sampling model.
- Want to find an alternative sampling density for the counts that is able to accommodate this additional variation.

Negative Binomial

- Recall y_i represents the number of "can"s in the i-th Federalist Papers.
- Conditional on parameters α and β, one assigns y_i the Negative Binomial density defined as

$$f(Y_i = y_i \mid \alpha, \beta) = \frac{\Gamma(y_i + \alpha)}{\Gamma(\alpha)} p_i^{\alpha} (1 - p_i)^{y_i},$$

where

$$p_i = \frac{\beta}{\beta + n_i/1000}.$$

Negative Binomial Generalizes Poisson

- Mean count is given by $E(y_i) = \mu_i$ where

$$\mu_i = \frac{n_i}{1000}\frac{\alpha}{\beta}.$$

- The ratio α/β is playing the same role as λ – one can regard α/β as the true rate of the particular word per 1000 words.
- The variance of the count y_j is given by

$$Var(y_i) = \mu_i \left(1 + \frac{n_i}{1000\beta}\right)$$

- With the extra multiplicative term $\left(1 + \frac{n_i}{1000\beta}\right)$, Negative Binomial is able to accommodate the additional variability in the counts.

Posterior Analysis

- Counts $y_1, ..., y_N$ are independent Negative Binomial with parameters α and β
- Likelihood function is equal to

$$L(\alpha, \beta) = \prod_{i=1}^{N} f(y_i \mid \alpha, \beta).$$

- Assume α and β are independent and assign to each α and β a Gamma density with parameters 0.001 and 0.001. Then the posterior density is given by

$$\pi(\alpha, \beta \mid y_1, \cdots, y_N) \propto L(\alpha, \beta)\pi(\alpha, \beta)$$

where $\pi(\alpha, \beta)$ is the product of Gamma densities.

R Work

- Using JAGS, the Negative Binomial density is represented by the JAGS function `dnegbin()` with parameters `p[i]` and `alpha`.
- One first defines `p[i]` in terms of the parameter `beta` and the sample size `n[i]`, and then expresses the Negative Binomial density in terms of `p[i]` and `alpha`.

```
modelString = "
model{
for(i in 1:N){
   p[i] <- beta / (beta + n[i] / 1000)
   y[i] ~ dnegbin(p[i], alpha)
}
mu <- alpha / beta
alpha ~ dgamma(.001, .001)
beta ~ dgamma(.001, .001)
}
```

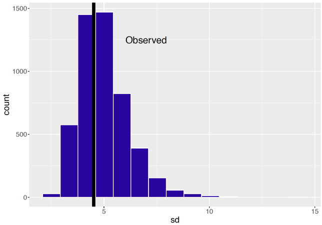
Posterior Predictive Checking

- Can the Negative Binomial density accommodate the extra variability in the word counts?
- One can check this statement by a posterior predictive check implemented in the R function `one_rep()`.
- Start with a simulated value (α^*, β^*) from the posterior distribution. Then we simulated a replicated dataset $y_1^R, ..., y_N^R$ where y_i^R has a Negative Binomial distribution with parameters α^* and $\beta^*/(\beta^* + n_i/1000)$. Then we compute the standard deviation of the $\{y_i^R\}$.

```
one_rep <- function(i){
  p <- post$beta[i] / (post$beta[i] + n / 1000)
  sd(rnbinom(length(y), size = post$alpha[i], prob = p)
}
```

Repeat Process

- By repeating this algorithm for 5000 iterations, store 5000 draws of the standard deviation of samples from the predictive distribution.
- Figure displays a histogram of the standard deviations from the predictive samples and the observed standard deviation of the counts is shown as a vertical line.
- Predictions with a Negative Binomial sampling model appear consistent with the spread in the observed word counts.



Inference

- One performs inferences about the mean use of the word "can" in Hamilton essays measured by the parameter $\mu = \alpha/\beta$.
- Figure displays MCMC diagnostic plots for the parameter μ. A 90% posterior interval estimate for the rate of "can" is (2.20, 3.29)