Chapter 13.2 Federalist Paper Study

Jim Albert and Monika Hu

Chapter 13 Case Studies

Introduction

- The Federalist Papers were a collection of articles written in the late 18th century by Alexander Hamilton, James Madison and John Jay to promote the ratification of the United States Constitution.
- Some of these papers were written by Hamilton, other papers were written by Madison, and the true authorship of some of the remaining papers has been in doubt.
- Mosteller and Wallace (1963) focused on the frequencies of counts of so-called filler words such as "an", "of", and "upon".
- The use of different sampling distributions is described to model word counts in a group of Federalist Papers.

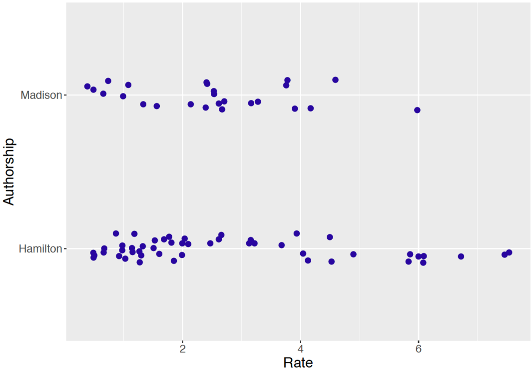
Data on word use

- Look at the occurrences of the word "can" in all of the Federalist Papers authored by Alexander Hamilton or James Madison.
- Table shows for each paper, the total number of words, the number of occurrences of the word "can" and the rate of this word per 1000 words.

	Name	Total	word	y	Rate	Authorship
1	Federalist No. 1	1622	can	3	1.85	Hamilton
2	Federalist No. 10	3008	can	4	1.33	Madison
3	Federalist No. 11	2511	can	5	1.99	Hamilton
4	Federalist No. 12	2171	can	2	0.92	Hamilton
5	Federalist No. 13	970	can	4	4.12	Hamilton
6	Federalist No. 14	2159	can	9	4.17	Madison

Graph

- Figure displays parallel jittered dotplots of the rates of "can" for the Madison and Hamilton papers.
- Note the substantial variability in the rates across papers.
- It appears that there is slight tendency for Hamilton to use this particular word more frequently than Madison.



Poisson density sampling

- Consider the word use of all of the Federalist Papers written by Hamilton.
- Initially assume that for the i-th paper the count y_i of the word "can" has a Poisson density with mean $n_i\lambda/1000$ where n_i is the total number of words and λ is the true rate of the word among 1000 words.

$$f(Y_i = y_i \mid \lambda) = \frac{(n_i\lambda/1000)^{y_i} \exp(-n_i\lambda/1000)}{y_i!}.$$

Likelihood

- Assuming independence of word use between papers, the likelihood function is the product of Poisson densities

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

- Posterior density of λ is given by

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

where $\pi()$ is the prior density.

Prior

- Suppose one knows little about the true rate of "can"s
- To reflect this lack of information, one assigns λ a Gamma density with parameters $\alpha = 0.001$ and $\beta = 0.001$.
- A JAGS script below is written to specify this Bayesian model. By use of the `run.jags()` function, one obtains a simulated sample of 5000 draws from the posterior distribution.

```
modelString = "
model{
for (i in 1:N) {
   y[i] ~ dpois(n[i] * lambda / 1000)
}
lambda ~ dgamma(0.001, 0.001)
}
"
```

Overdispersion?

- With count data, one general concern is **overdispersion**
- Do the observed counts display more variability than one would anticipate with the use of this Poisson sampling model?
- One can check for overdispersion by use of a posterior predictive check.

Posterior predictive checking

- One simulates one replicated dataset from the posterior predictive distribution as follows.

1. One simulates a value of λ from the posterior distribution.
2. Given the simulated value $\lambda = \lambda^*$, one simulates counts $y_1^R, ..., y_N^R$ from independent Poisson distribution with means $n_1\lambda^*/1000, ..., n_N\lambda^*/1000$.

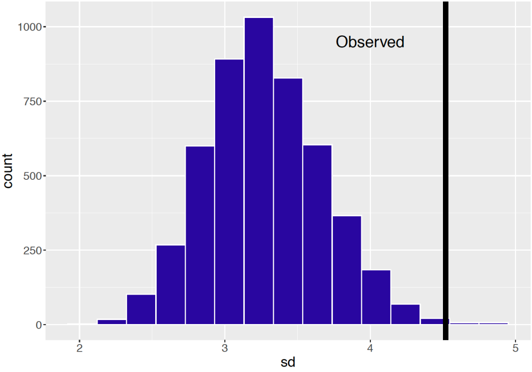
Checking function

- Given a replicated dataset of counts $\{y_i^R\}$, one computes the standard deviation.
- A standard deviation is a reasonable choice of a testing function since one is concerned about the variation in the data.
- R script implements simulation.

```
one_rep <- function(i){
  lambda <- post[i]
  sd(rpois(length(y), n * lambda / 1000))
}
sapply(1:5000, one_rep) -> SD
```

Repeat

- Repeat this process 5000 times, obtaining 5000 values of the standard deviation.
- Figure displays a histogram of the standard deviations from the predictive distribution and the standard deviation of the observed counts $\{y_i\}$ is displayed as a vertical line.



Interpret

- The observed standard deviation is very large relative to the standard deviations of the counts from the predictive distribution.
- We see evidence of overdispersion.
- There is more variability in the observed counts of "can"'s than one would predict from the Poisson sampling model.