Project: Bayesian Statistics

Simon

12/6/2017

**Fitted with Logistics Regression**

O/E: 0.9935365

Sd for O/E: 0.3575793

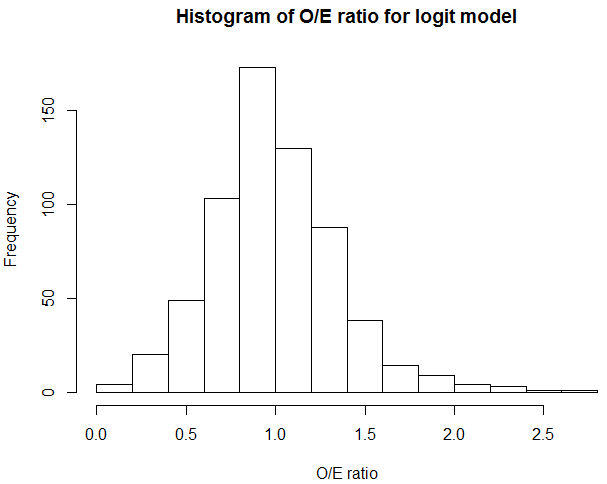
95% CI for OE: (0.5978625 2.5900719)

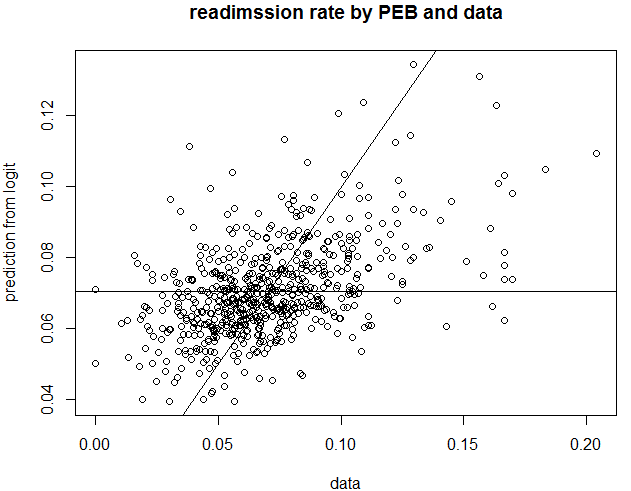
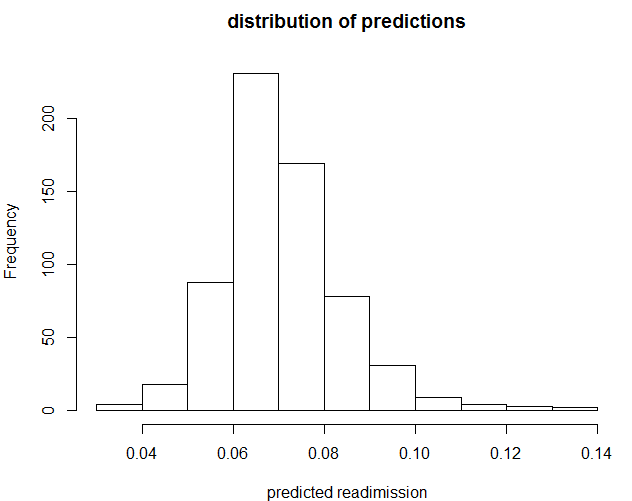
Mean for Readmission rate:0.07129507

Median for Readmission rate: 0.06891252

Variance: 0.0001766114

Standard deviation: 0.01328952





**Level 2 Model**

OtoE

-------------------------------------------------------------

Percentiles Smallest

1% .8294251 .7778019

5% .8780756 .782155

10% .9118117 .7951289 Obs 637

25% .9533957 .7963859 Sum of Wgt. 637

50% .9933739 Mean 1.002794

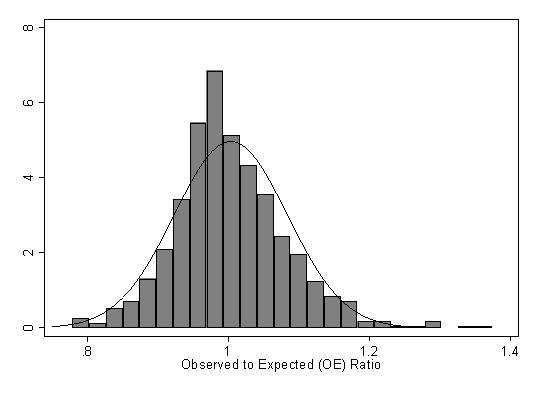
Largest Std. Dev. .0803928

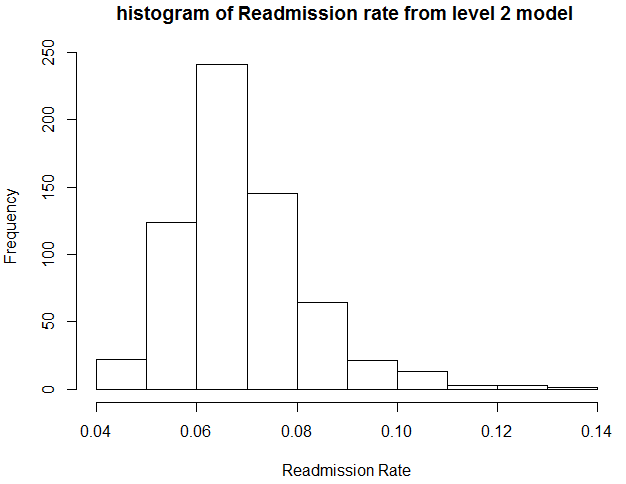
75% 1.04606 1.287822

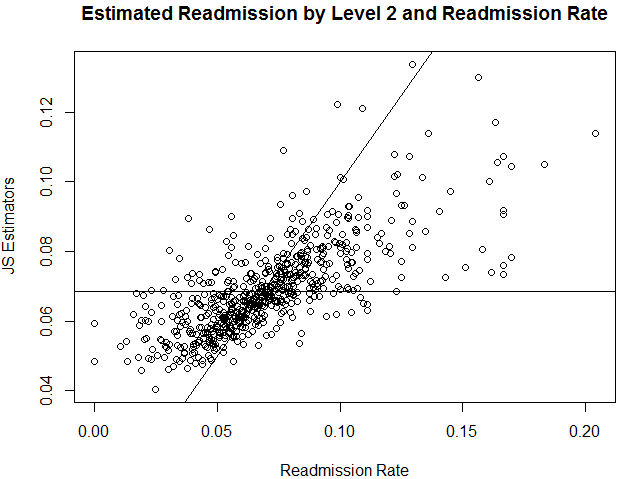
90% 1.102034 1.294576 Variance .006463

95% 1.143232 1.344361 Skewness .6262953

99% 1.236607 1.373304 Kurtosis 4.611423







**Empirical Bayes (EB)**

**Parametric EB**

O/E: 1.031609

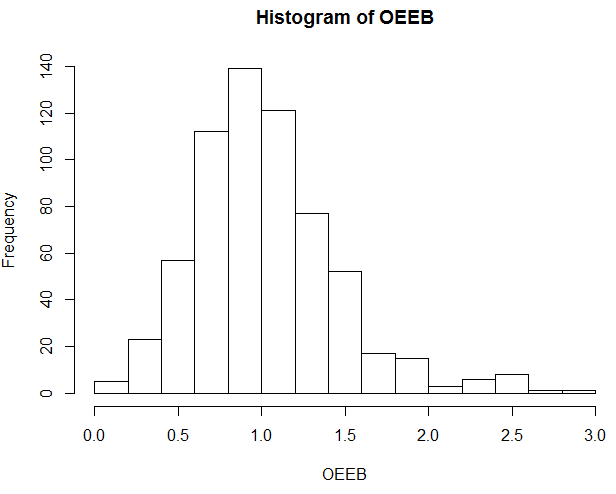
Sd for O/E: 0.4279435

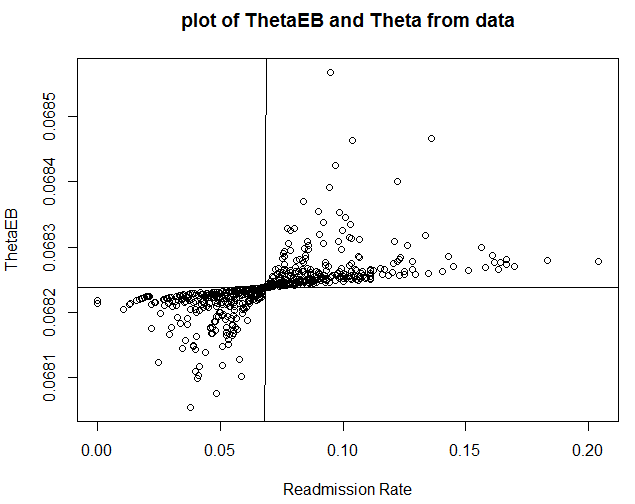
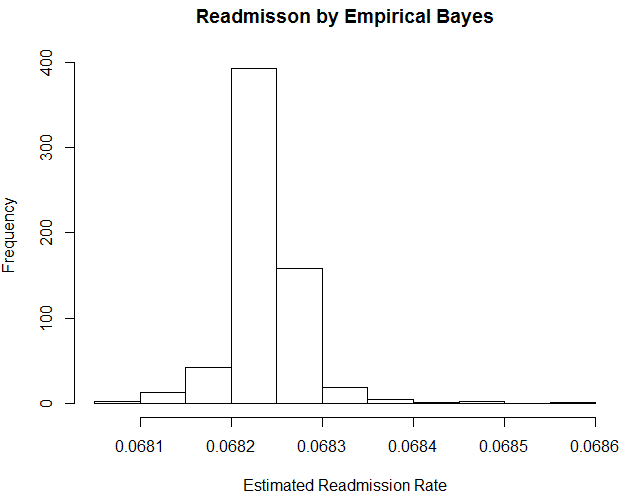
95% CI for OE: (0.3180131 2.1316855)

Mean for Readmission rate: 0.06823827

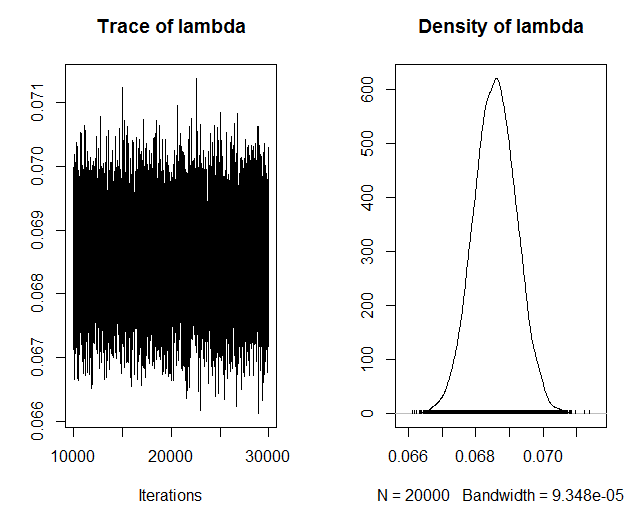
Median for Readmission rate: 0.06823681

Standard deviation: 4.067923e-05





**MCMC: Poisson Gamma Model**



1. Empirical mean and standard deviation for each variable,

plus standard error of the mean:

Mean SD Naive SE Time-series SE

6.857e-02 6.392e-04 4.520e-06 4.520e-06

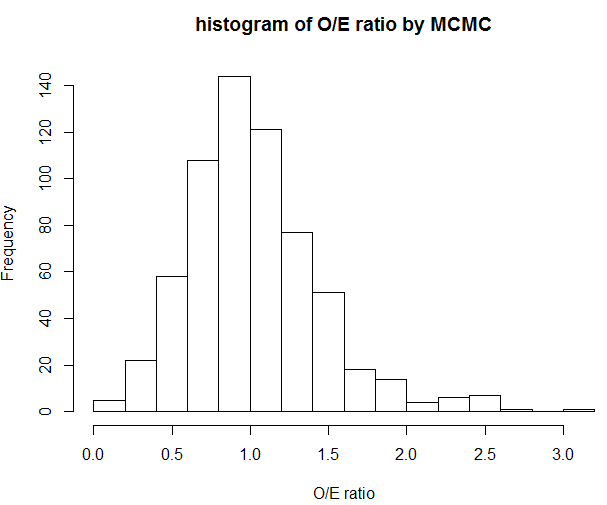
2. Quantiles for each variable:

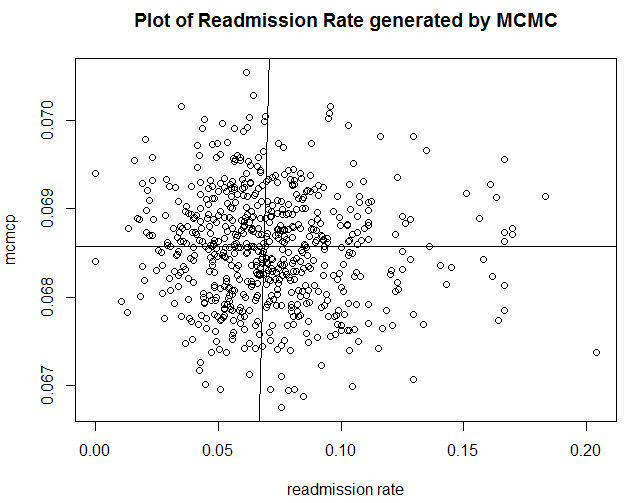
2.5% 25% 50% 75% 97.5%

0.06731 0.06814 0.06858 0.06900 0.06983

3. O/E ratio: 1.027538

95% CI for O/E ratio: (0.3155265 2.1270785)





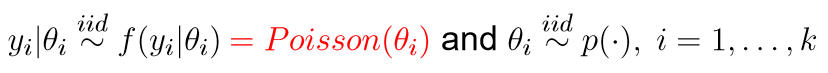
**Nonparametric EB**

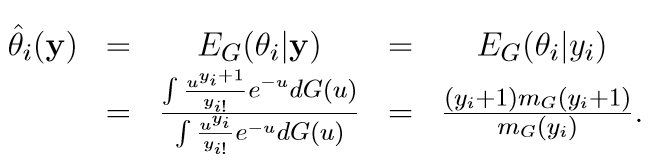
Introduction

Orbanz and Teh ()“A Bayesian nonparametric model is a Bayesian model on an inﬁnite-dimensional parameter space. The parameter space is typically chosen as the set of all possible solutions for a given learning problem.”

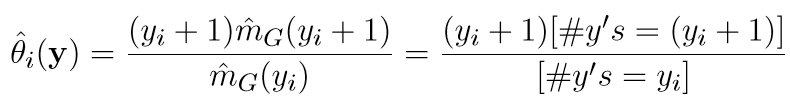
Method:

We Denote yi are numbers of readimission of hospitals from data, θi are predicted numbers of readmission for each hospital, then assume only that the θi are i.i.d from some distribution p. We use the data to estimate the prior or the marginal distribution directly.





And Robbins Miracle estimated each θi directly:



Result

mean for O/E: 0.7392372

Sd for O/E: 0.1209511

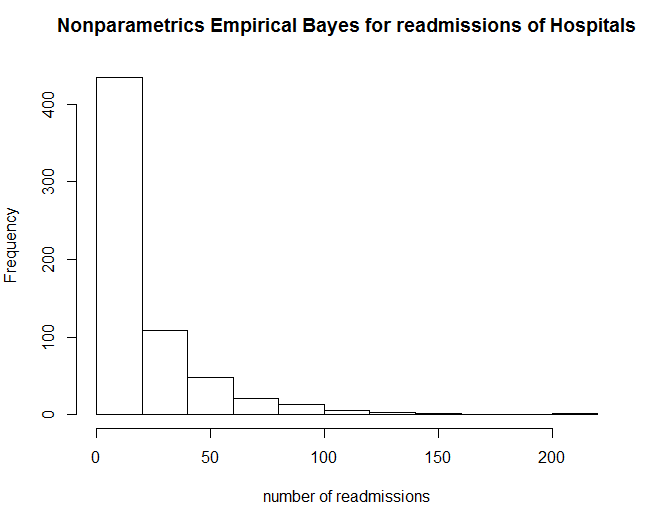
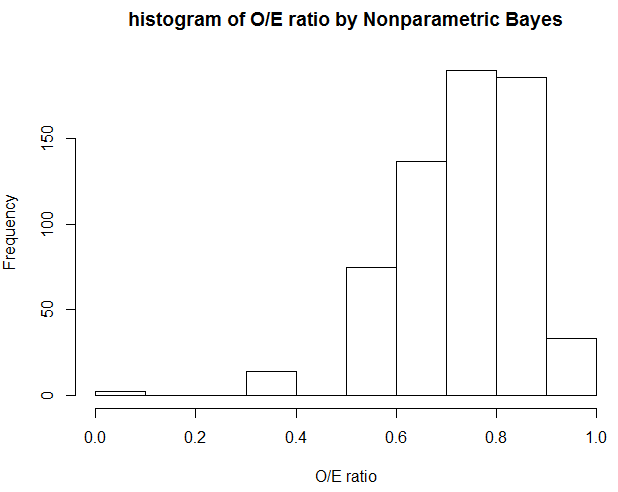
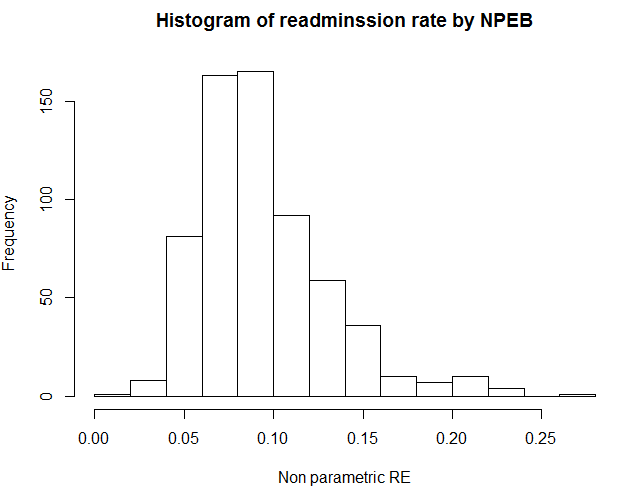
95% CI for OE: (0.5136842 0.9118940)

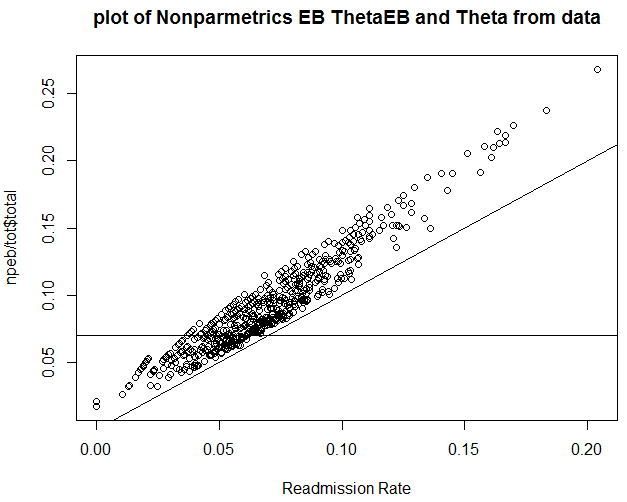
Mean for Readmission rate: 0.09444687

Median for Readmission rate: 0.08698843

Variance: 0.001341436

Standard deviation: 0.03662562



Disscussion:

Reference:

L. Wasserman. All of Nonparametric Statistics. Springer, 2006.

Peter Orbanz, Cambridge University, Yee Whye Teh, University College London, Bayesian Nonparametric Models

**James-stein Estimator**

Introduction

It is similar for sports that they are applied with statistics. Efron and Morris investigated the performance of baseball player. They used Bayesian Analysis that they collected the data of the previous performance and the seasonal performance. By using the Bayesian statistics, and “The best guess about the future is usually obtained by computing the average of past events”. They recorded bat times, hit times and calculate the hit ratio for each player yi, grand average for hit ratio y bar, which are between 0 and 1 (proportion is hit times/ bat times). Then Efron and Morris focused on the important part of the James-Stein’ estimator: z=y bar+ c \* (yi-y bar). Where c is the degree of shrinkage. This shows the “shrinkage” that the differences between each player’s performance and the overall average performance.

z=y bar+ c \* (yi-y bar)

Method

As discussed before, the James-Stein’s Estimator is

,

The c is treated as a constant, but here it is depended by the sample but not a constant,



And k is number of unknown mean.

Result:

O/E: 0.9744177

Sd for O/E: 0.2677618

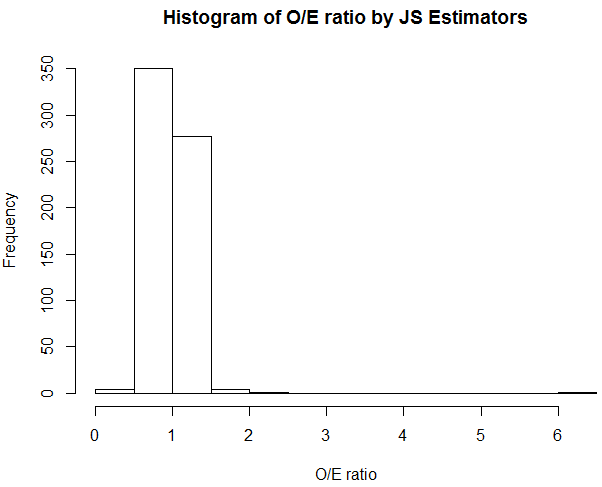
95% CI for OE: (0.6038826 1.2682973 )

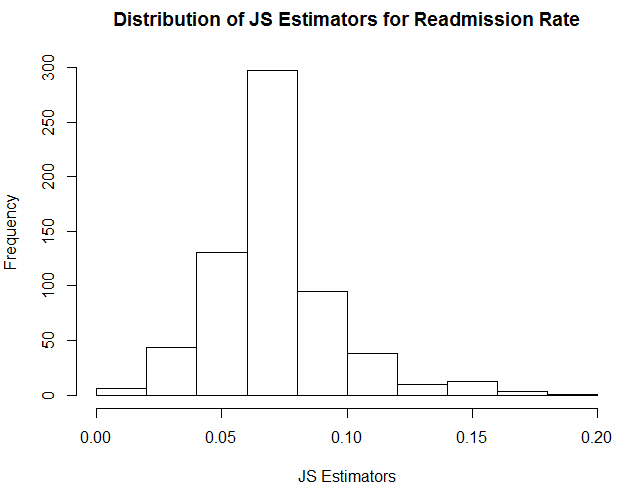
Mean for Readmission rate:0.07129507

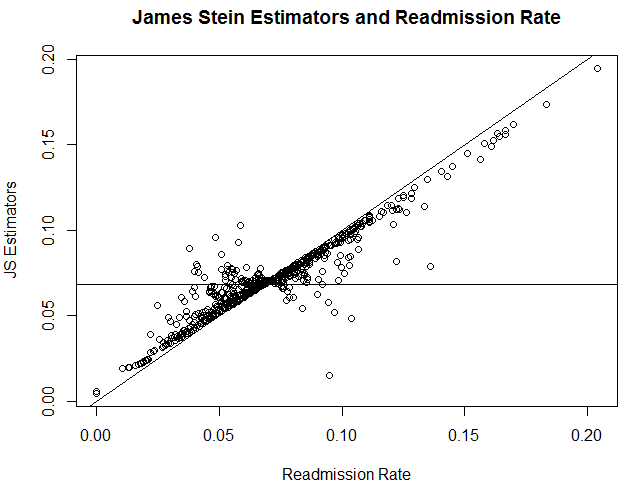
Median for Readmission rate: 0.06985986

Variance: 0.0006055711

Standard deviation: 0.02460835





Discussion:

Comparing to the result of Empirical Bayes, this shows that the MSE is larger than MSE generated by MLE.