Evidence to use Negative Binomial:

Dispersion = 3.118991, should be around 1 if not overly dispersed

Overdispersion test

data: readmission\_model

z = 33.347, p-value < 2.2e-16

alternative hypothesis: true dispersion is greater than 1

sample estimates:

dispersion

3.867938

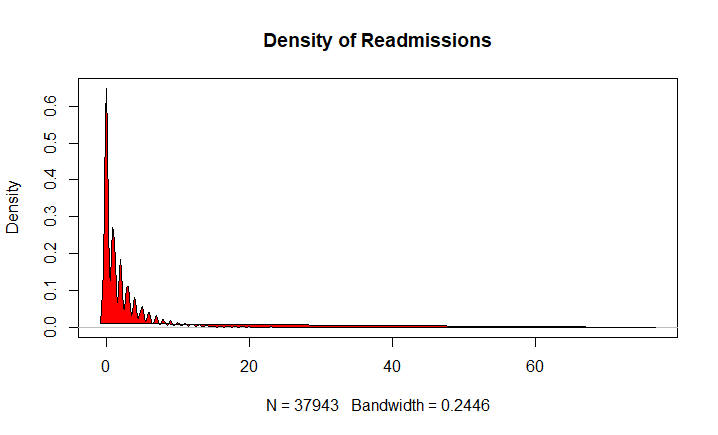
AIC Comparison:

Poisson: 183082.5

Neg Binomial: 136241.1

Akaike Information criterion (AIC) is insanely lower for Negative Binomial. Much better fitting model.

Zero-Inflated:



High Density around zero value supports using zero-inflated model. Also the wide distribution supports using negative binomial .

AIC Comparison:

Negative Binomial: 136241.1

Zero-Inflated Negative Binomial: 135766.1

Zero-inflated Negative binomial is lower.

Vuong Test:

Vuong Test is used to compare non-nested count models, such as negative binomial vs zero-inflated negative binomial. Will adjust for differences in the models and provide more adjusted comparison of AIC.

Result:

Vuong Non-Nested Hypothesis Test-Statistic:

(test-statistic is asymptotically distributed N(0,1) under the

null that the models are indistinguishible)

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

Vuong z-statistic H\_A p-value

Raw -10.681571 model2 > model1 < 2.22e-16

AIC-corrected -9.738565 model2 > model1 < 2.22e-16

BIC-corrected -5.710118 model2 > model1 5.6449e-09

Significant result supports that Zero-inflated Negative Binomial Model (Model 2) better fits the data than the Negative Binomial Model (Model 1) for both the Akaike Information criterion (AIC) and Bayesian Information Criterion (BIC).

Conclusion:

All of this supports that the distribution for our count outcome, 30-day readmissions, is overly dispersed and highly inflated at zero, indicating that a zero-inflated negative binomial model should be used.