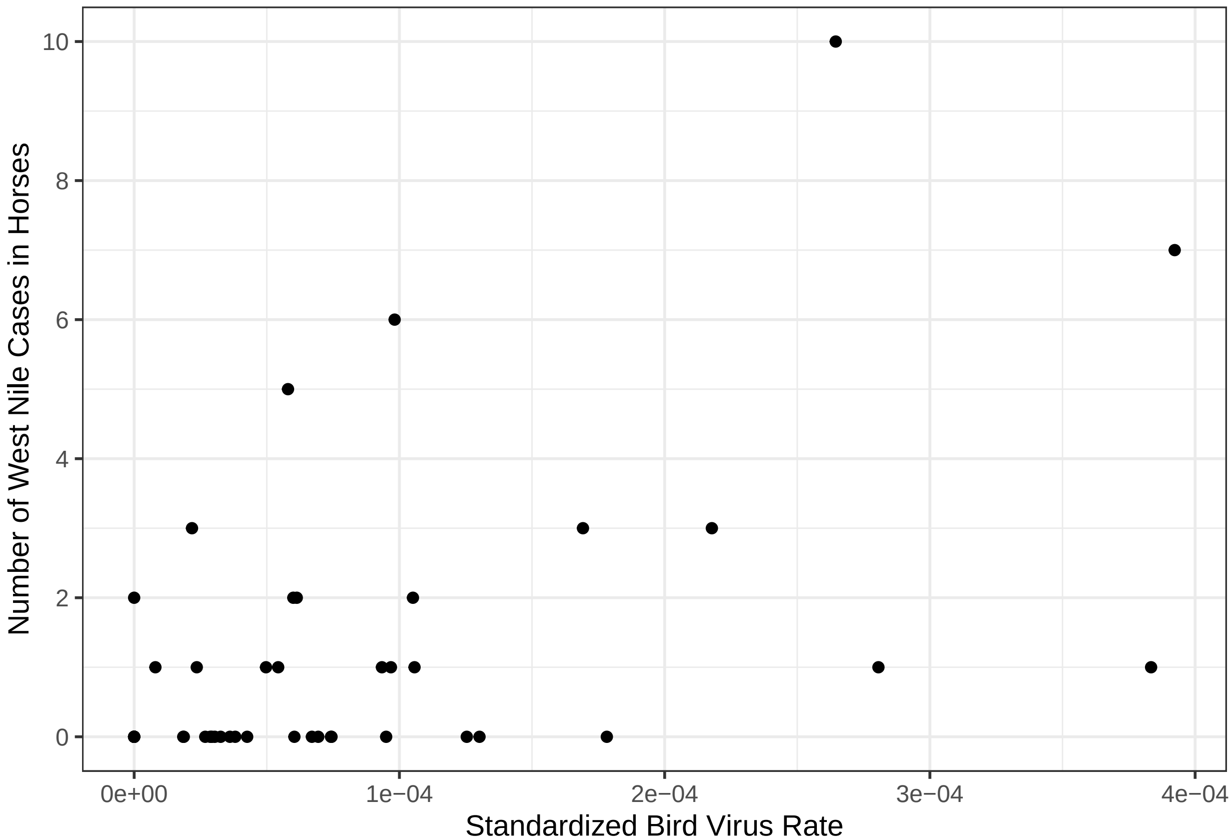
2018 Qual - Problem 3

Elaine Kaye Kearney

6/27/2020

2018 Qual - Problem 3

a)


Number of West Nile Cases in Horses
Standardized Bird Virus Rate
0e+00
1e−04
2e−04
3e−04
4e−04
0
2
4
6
8
10

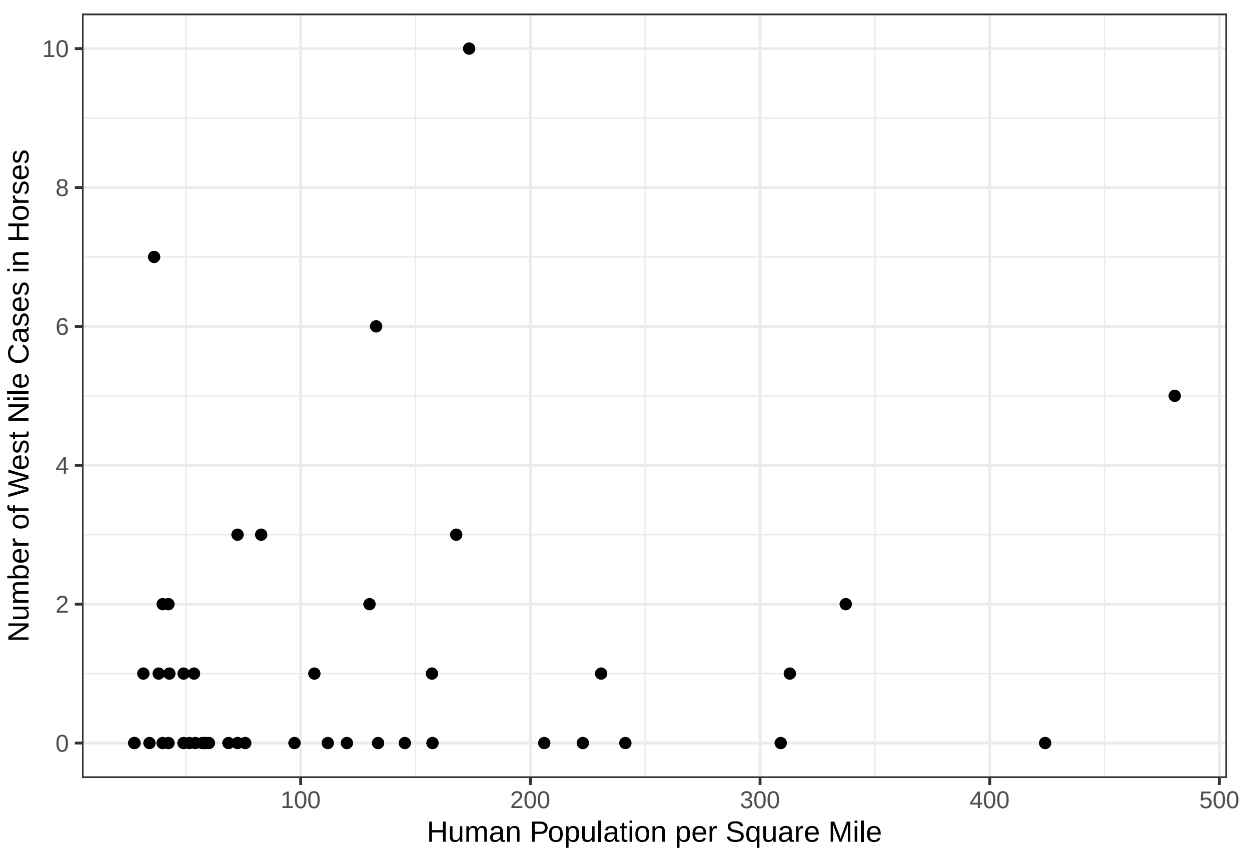

Number of West Nile Cases in Horses
Human Population per Square Mile
100
200
300
400
500
0
2
4
6
8
10

Table 1: Summary Statistics for Number of West Nile Cases in Horses

N	Number Missing	Mean	Standard Deviation	Minimum	Maximum
46	0	1.17	2.101	0	10

b)

Let μ_{ij} be the expected number of cases of West Nile virus in horses per farm for the ith standardized bird virus rate and jth human population density.

$$log(\mu_{ij}) = \alpha + log(N_{ij}) + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \beta_3 X_{ij1} X_{ij2}$$

where X_{ij1} is the standardized bird virus rate and X_{ij2} is the human population density. α = log expected rate of cases of West Nile virus in horses per farm for standardized bird virus rate of 0 and human population density of 0. $log(Nij)$ = log of exposure for the ith standardized birth virus rate and jth human population density (the offset). β_1 = increment in log expected rate of cases of West Nile virus in horses per farm for each one unit increase in standardized bird virus rate. β_2 = increment in log expected rate of cases of West Nile virus in horses per farm for each one unit increase in human population density. β_3 = interaction term for pairwise interaction between standardized bird virus rate and human population density.

Table 2: Parameter Estimates and Standard Error Estimates for Poisson Model of Standardized Horse Virus Rate

Parameter	Parameter Estimate	Standard Error Estimate
α	-6.98	0.366
β_1	5329.84	1595.642
β_2	0.0006	0.0017
β_3	26.01	12.382

I used a Wald Chi-square test to test the null hypothesis of no association between West Nile virus in horses and West Nile virus in birds, which is equivalently the test that $\beta_1 = \beta_3 = 0$. This test produced a Wald Chi-square statistic of 53.13, with 2 degrees of freedom, which corresponds to a p-value of < 0.0001. Therefore, we have strong evidence to reject the null hypothesis of no association between West Nile virus in horses and West Nile virus in birds.

c)

For the model in b, we assumed that the count of West Nile virus in horses follows a Poisson model. We know that for a Possion distribution, the mean is equal to the variance. Concern about possible overdispersion is that the varaince of the counts of West Nile virus in horses is larger than the mean of the counts of West Nile virus in horses.

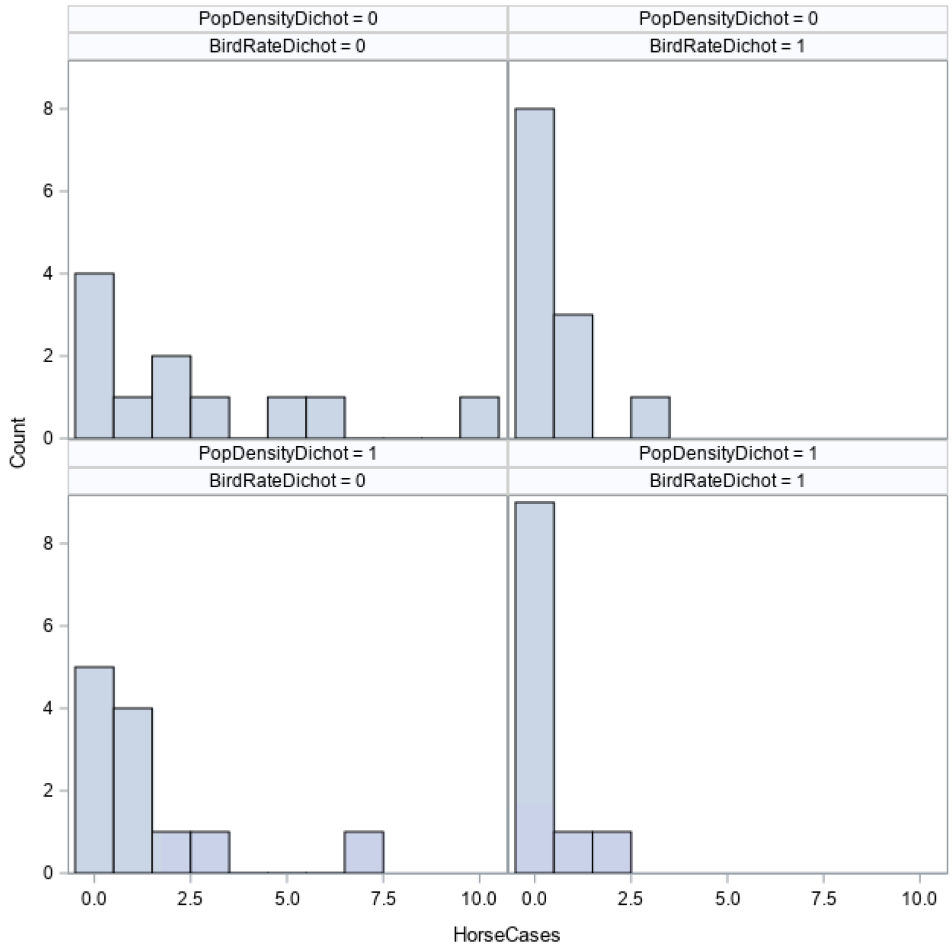
Table 3: Parameter Estimates and Standard Error Estimates for Negative Binomial Model of Standardized Horse Virus Rate

Parameter	Parameter Estimate	Standard Error Estimate
α	-6.92	0.423
β_1	4516.53	2474.606
β_2	-0.0007	0.0024
β_3	44.64	28.516

I used a Wald Chi-square test to test the null hypothesis of no association between West Nile virus in horses and West Nile virus in birds, which is equivalently the test that $\beta_1 = \beta_3 = 0$. This test produced a Wald Chi-square statistic of 20.40, with 2 degrees of freedom, which corresponds to a p-value of < 0.0001. Therefore, we have strong evidence to reject the null hypothesis of no association between West Nile virus in horses and West Nile virus in birds.

d)

With excess zeroes in the data, when we fit the Poisson model we may be underestimating the Poisson parameter, which is the mean, because the mean is drawn down smaller because of the excess zeroes in the data. In order to address this concern, I will refit the Poisson model using exact Poisson regression.



Poisson Model:

Table 4: Parameter Estimates and Standard Error Estimates for Poisson Model of Standardized Horse Virus Rate Assuming Zero-Inflated Poisson

Parameter	Parameter Estimate	Standard Error Estimate
α	-6.37	0.483
β_1	3845.25	1854.999
β_2	0.0001	0.0019
β_3	21.42	13.230

Zero Model:

Table 5: Parameter Estimates and Standard Error Estimates for Zero Model of Standardized Horse Virus Rate Assuming Zero-Inflated Poisson

Parameter	Parameter Estimate	Standard Error Estimate
α	0.49	1.476
β_1	-19699.2	33996.24
β_2	0.005	0.0097
β_3	-147.84	230.061

Tested the null hypothesis of no overdispersion and resulted in Scaled Pearson χ^2 of 57.6537 with df = 38, which corresponds to a p-value of 0.02. Therefore, there is evidence of overdispersion which means inferences based on these estimates are suspect and the standard errors are likely to be biased downwards.

e)

Negative Binomial Model:

Table 6: Parameter Estimates and Standard Error Estimates for Negative Binomial Model of Standardized Horse Virus Rate Assuming Zero-Inflated Negative Binomial

Parameter	Parameter Estimate	Standard Error Estimate
α	-6.38	0.515
β_1	3604.13	2173.857
β_2	-0.0002	0.0022
β_3	26.63	20.091

Zero Model:

Table 7: Parameter Estimates and Standard Error Estimates for Negative Binomial Model of Standardized Horse Virus Rate Assuming Zero-Inflated Negative Binomial

Parameter	Parameter Estimate	Standard Error Estimate
α	0.44	1.552
β_1	-21907.7	36141.74
β_2	0.004	0.0102
β_3	-140.51	242.665

Tested the null hypothesis of no overdispersion and resulted in Scaled Pearson χ^2 of 51.8117 with df = 38, which corresponds to a p-value of 0.07. Therefore, we fail to reject the null hypothesis of no overdispersion.

But may only need bird rate in this model.