Milestone 3

Andres Calderon
acald013@ucr.edu

February 26, 2016

1 Formal reasoning.

Figure Xa shows the difference in the accuracy results of kNN when the probability of error in the distance computation increase. Figure Xb show the error rate of the same implementation by just substracting the value of the original accuracy (without error injection) and the values obtained in figure Xa. We can see that the increasing rate of error described by figure Xb can be modeled by an exponential function using the equation $Err(p) = \alpha \times e^{\beta \times p}$.

Equation 1 illustrates the result.

$$Err(p) = (1 - dr) \times e^{\frac{N}{C} \times (1 - dr) \times p} \tag{1}$$

2 Testing.

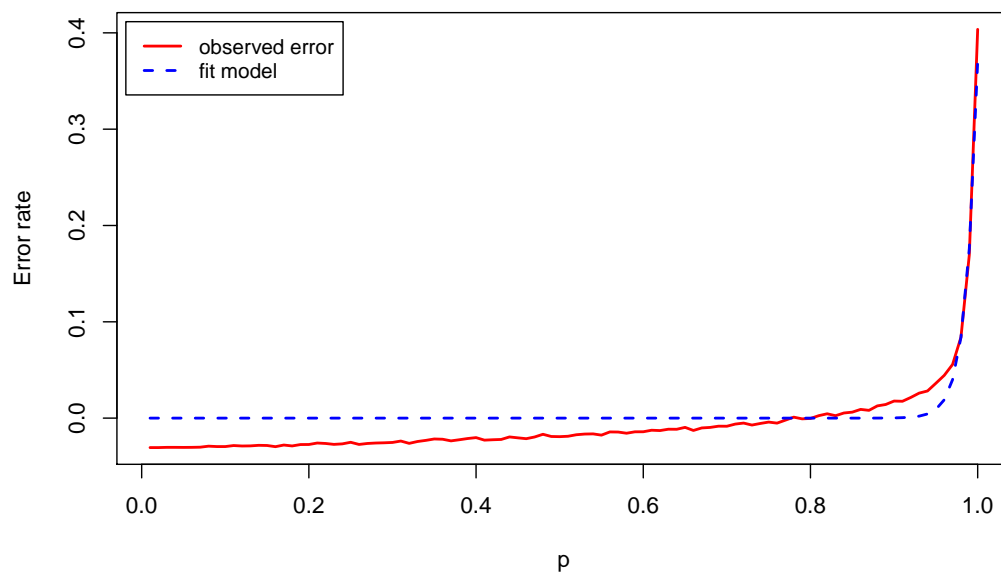|	dr	N	C	D
Cancer	0.63	398	2	32
Iris	0.33	105	3	4
Seeds	0.33	147	3	7
Wine	0.40	125	3	13
Zoo	0.41	71	7	17

Table 1: Description of the datasets.

Figure 1: Observed error rate compared to the fitted model for Cancer.
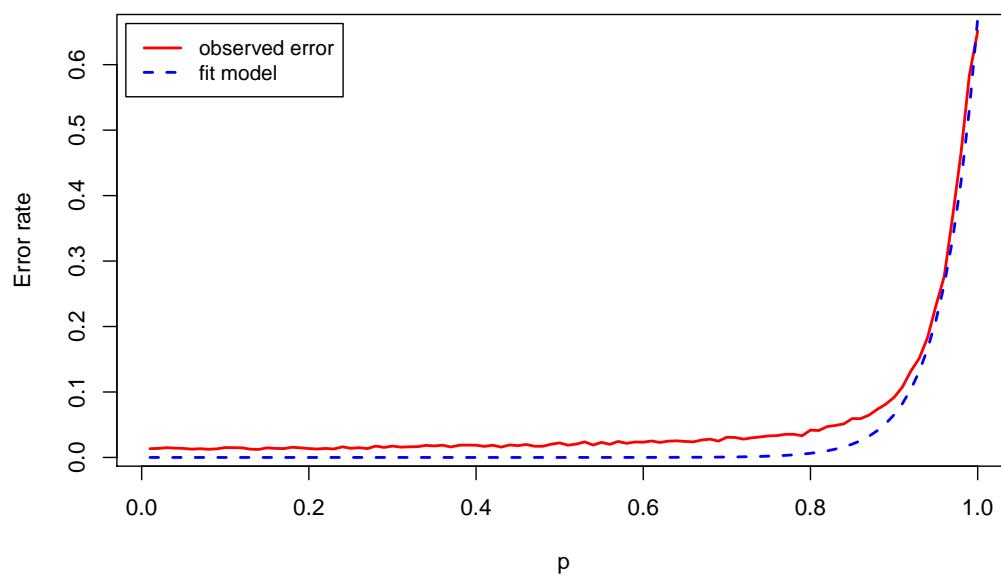

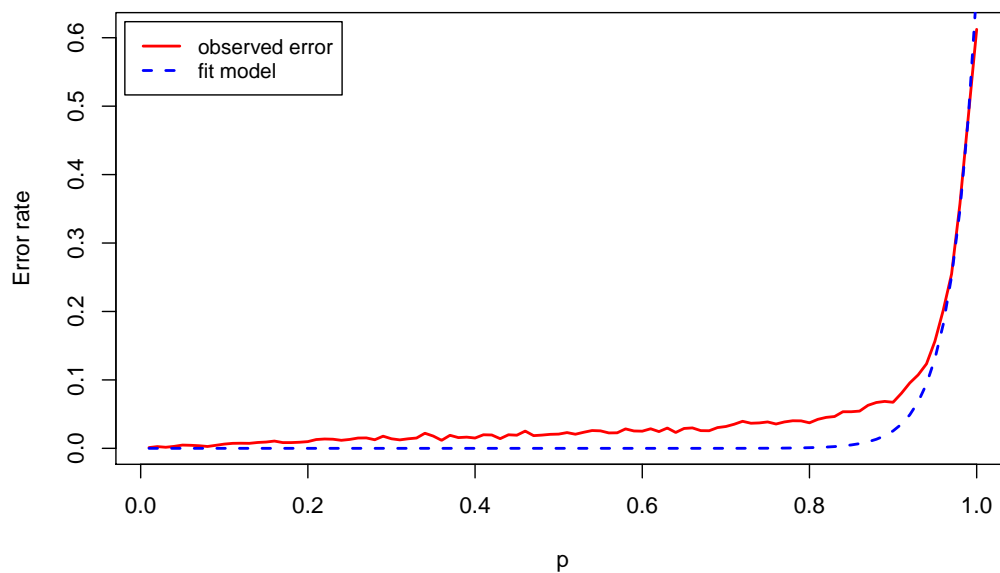Figure 2: Observed error rate compared to the fitted model for Iris.

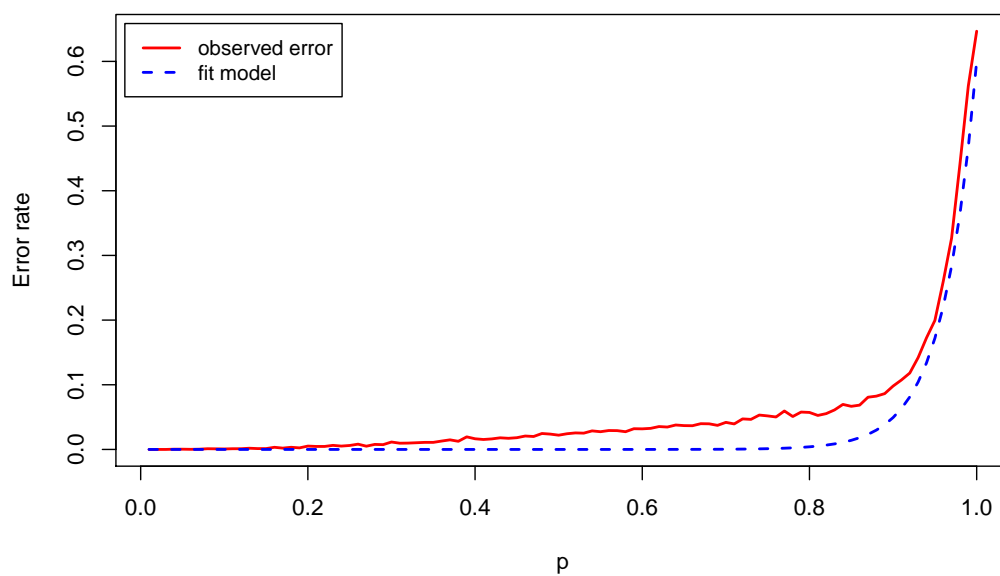Figure 3: Observed error rate compared to the fitted model for Seeds.

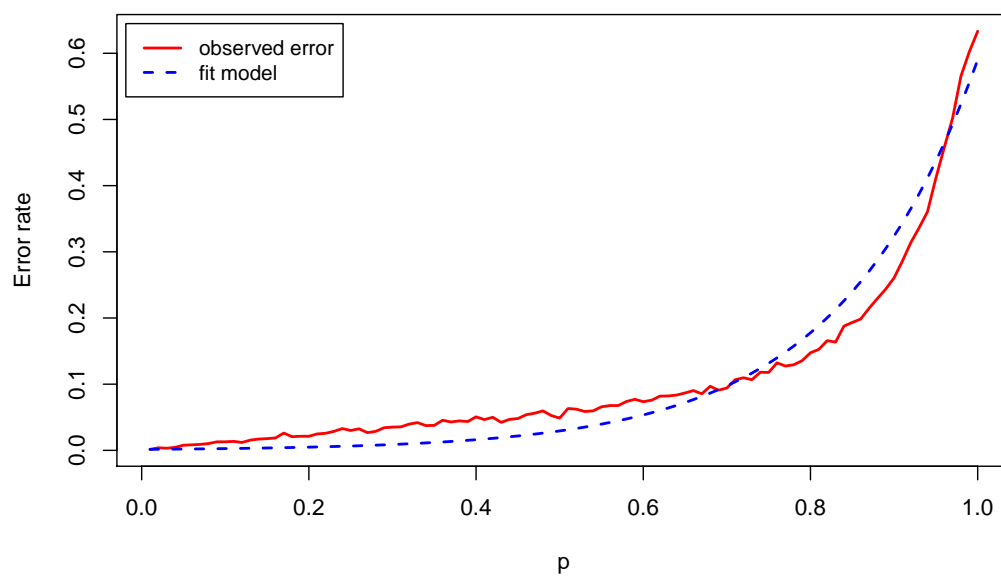Figure 4: Observed error rate compared to the fitted model for Wine.

Figure 5: Observed error rate compared to the fitted model for Zoo.