1 Probability decision boundary

We want to predict y = 1 if the expected loss of guessing y = 1 is less than the expected loss of guessing y = 0. P(y = 0 / x) * 5 = P(y = 1 / x) * 10. Let p1 = P(y = 1 / x). Then, (1 - p1) * 5 <p1 * 10, so p1 >1/3. The threshold should be set to 1/3.

2 Double counting the evidence

(a) If we consider only X_1 we get an error rate of $0.1 + 0.15 = 0.25$
if we use only X_2 and the error rate is $0.25 + 0.05 = 0.30$
(b) The error rate of the Naive Bayes classifier is given by

$$error = \sum_{X_1, X_2, Y} [f(X_1, X_2) \neq Y] \Pr(X_1, X_2, Y)$$

To compute the error rate using X_1 and X_2 we sum $\Pr(X_1, X_2, Y)$ for all the events where $f(X_1, X_2) \neq Y$, we get $0.05 + 0.035 + 0.135 + 0.015 = 0.235$, which is better than if using only a single attribute (X1 or X2).

(c) The error rate is obtained by summing the probability of all events that make wrong predict for Y which is $0.05 + 0.035 + 0.20 + 0.015 = 0.3$.

(d) Naive Bayes is based on the assumptions of conditional independence. Naive Bayes performs worse because adding a X3 which is dependent of X2 doesn't introduce any additional information.

(e) Where as Logistic Regression does not have the same conditional independence assumptions and so it does not suffer from introducing X3 which is the exact copy of X2.