Generative Models for Classification

In the case of Logistic Regression we directly model $\Pr(Y = k|X = x)$ using the Sigmoid Function (Logistic Function) its a simple conditional probability approach **Predicting Y given X**.
Considering the alternative Probability of the the the predictors X given a class Y which is the logic behind **Bayes Theorem**, When the distribution of X is normal the model turn to be very similar to Logistic Regression.

Why its needed?

- If There is a substantial separation between the two classes the **parameter estimates** for logistic regression model will be unstable
- If the predictors X are normally distributed and the sample size n is small Using **Generative Models** will be more accurate than Logistic Regression

Suppose that we want to classify an observation into on of the K classes where$K \geq 2$, means the Response Y can take K possible classes.

- Let π_k be the **Prior Probability** that the a random Observation comes from the $k\text{th}$ class
- Let $f_k(X) = \Pr(X|Y = k)$ is the **Density Function** of X
  - $f_k(x)$ will be large if there is **high probability** of that observation being in the $k\text{th}$ class

$$\Pr(Y = k|X = x) = \frac{\pi_k f_k(x)}{\sum^K \pi_i f_i(x)}$$

- This is just the Bayes' theorem Formula
- Let $p_k(x) = \Pr(Y = k|X = x)$ and will also be called the **Posterior** probability
- So our goal is to know the probability of the Observation being in the $k\text{th}$ class *Given* the class k aka Response Y