1. Logistic posterior with exponential class conditionals

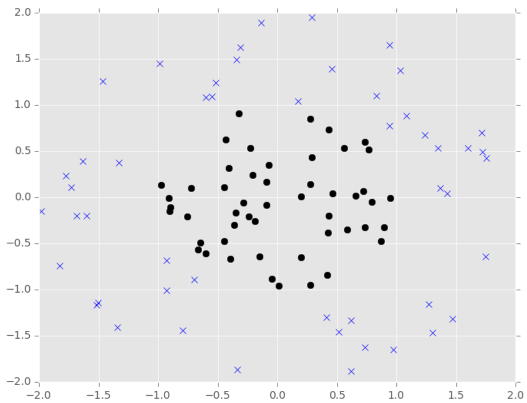
   We have seen in class that Gaussian class conditionals lead to a logistic posterior that is quadratic in X. Now, suppose the class conditionals are exponentially distributed with paramateres λ_i, i.e.

$$X|Y = i \sim \lambda_i \exp(-\lambda_i x), \quad \text{where } i \in \{0, 1\}$$
$$Y \sim \text{Bernoulli}(\pi)$$

   Show that the posterior distribution of the class label given X is also a logistic function, however with a linear argument in X. Assuming 0-1 loss, what will the decision boundary look like (i.e., describe what the posterior probability plot looks like)?

2. Circular Distributions

   Consider the following dataset



   where each point $x_n = (x_{1,n}, x_{2,n})$ is sampled *iid* and uniformly at random from two equiprobable (each equally likely) classes, a disk of radius 1 ($y_n = 1$) and a ring from 1 to 2 ($y_n = -1$).

   Qualitatively estimate and sketch the following quantities

   - The class conditional densities $f_{X|Y=c_i}$
   - The density of X
   - The posterior probabilities $P\left(Y = c_i|X\right)$
   - The Bayes Classifier
   - The Bayes Risk

3. Bayesian Decision Theory: Case Study - We're going fishing!

We want to design an automated fishing system that captures fish, classifies them, and sends them off to two different companies, Salmonites, Inc., and Seabass, Inc. For some reason we only ever catch salmon ($Y = 1$) and seabass ($Y = 2$). Salmonites, Inc. wants salmon, and Seabass, Inc. wants seabass. Given only the weights of the fish we catch, we want to figure out what type of fish it is using machine learning!

Let us assume that the weight of both seabass and salmon are both normally distributed (univariate Gaussian), given by the p.d.f.:

$$P(x|Y = i) = \frac{1}{\sigma_i\sqrt{2\pi}}e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}$$

We are given this data:

Data for salmon: $\{3, 4, 5, 6, 7\}$

Data for seabass: $\{5, 6, 7, 8, 9, 7 + \sqrt{2}, 7 - \sqrt{2}\}$

When we classify seabass incorrectly, it gets sent to Salmonites, Inc. who won't pay us for the wrong fish and sells it themselves. When we classify salmon incorrectly, it gets sent to SeaBass, Inc., who is nice and returns our fish. This situation gives rise to this loss matrix:

Predicted:

Truth:		**salmon**	**seabass**
	salmon	0	1
	seabass	2	0

a) First, compute the sample mean $\hat{\mu}_i$ and variance $\hat{\sigma}_i^2$ for the univariate Gaussian in both the seabass and the salmon case. Also compute the empirical estimates of the priors $\hat{\pi}_i$.

$\hat{\mu}_1 =$ $\hat{\mu}_2 =$

$\hat{\sigma}_1^2 =$ $\hat{\sigma}_2^2 =$

$\hat{\pi}_1 =$ $\hat{\pi}_2 =$

What is significant about $\hat{\sigma_1}$ and $\hat{\sigma_2}$?

b) Next, find the decision rule when assuming a 0-1 loss function. Recall that a decision rule for the 0-1 loss function will minimize the probability of error.

c) Now, find the decision rule using the loss matrix above. Recall that a decision rule, in general, minimizes the risk, or expected loss.