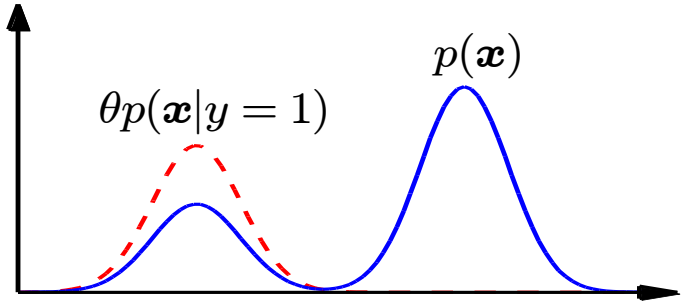
T-32 Learning from Positive and Unlabeled Data 2: Computationally Efficient Estimation of Class Priors

Marthinus C. du Plessis (UTokyo), Gang Niu (Baidu Inc.) and Masashi Sugiyama (UTokyo)

Task: Estimate the class prior π from the *positive and unlabeled* data

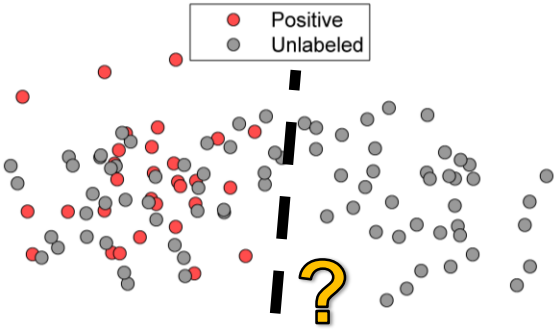
We show that π can be estimated by partially matching two distributions



Overestimation is avoided by penalization, giving a simple estimator

Why is π needed?

To train a classifier we need to know π (**T-31**)



$$\mathcal{X} := \{\boldsymbol{x}_i\}_{i=1}^{n} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x}|y=1)$$

$$\mathcal{X}' := \{\boldsymbol{x}'_i\}_{i=1}^{n'} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x})$$

$$p(\boldsymbol{x}) = \pi p(\boldsymbol{x}|y=1) + (1-\pi)p(\boldsymbol{x}|y=-1)$$

$$\pi = p(y=1) \quad \textbf{\textit{Class prior?}}$$