Estimating divergence functionals and the likelihood ratio by penalized convex risk minimization

- given n resp. m i.i.d. samples:
$$\begin{array}{rcl} X_1, \ldots X_n & \sim & \mathbb{P} \text{ (unknown)} \\ Y_1, \ldots Y_m & \sim & \mathbb{Q} \text{ (unknown)} \end{array}$$

- propose, analyze and evaluate convergence behavior of a nonparametric estimator for divergence functional $D_\phi(\mathbb{P}, \mathbb{Q})$ based on (X, Y):

$$D_\phi(\mathbb{P}, \mathbb{Q}) := \int \phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) d\mathbb{P} = \int \phi\left(\frac{q_0}{p_0}\right) p_0 d\mu.$$

  - also estimator of likelihood ratio function $\frac{d\mathbb{P}}{d\mathbb{Q}}$

- applications: estimating mutual information and KL divergence in coding, independent component analysis, dimensionality reduction, experiment design,...

- idea: based on variational characterization of divergence functionals