Predicting the label Y of an object X is a core task in machine learning. From a probabilistic perspective, this involves reasoning about conditional probabilities $p(y|x)$. However, it is hard to obtain reliable estimates for these probabilities. Here we show how to obtain lower and upper bounds on $p(y|x)$ given statistical information, and show how it can be used within various learning setups. We also extend this formulation to the structured case, where $y$ can be multivariate.