Problem 2.1

Assume that you have to predict a sequence $Y_1, Y_2, \ldots \in \{0, 1\}$ of i.i.d. random variables with unknown distribution. Your decision space is $[0, 1]$, and the loss function is

$$\mathcal{L}(\hat{P}, y) = |\hat{P} - y|.$$

How would you proceed? Try to estimate the cumulative loss of your forecaster and compare it to the cumulative loss of the best of the two experts, one of which always predicts 1 and the other always predicts 0.

Which are the most "difficult" distributions? How does your (expected) regret compare to that of the weighted average algorithm (which does not "know" that the outcome sequence is i.i.d.)?